

UNIVERSIDAD DE CONCEPCIÓN

Facultad de ingeniería

Departamento de Ingeniería Metalúrgica

Profesor Patrocinante

Sr. Roberto Fustos

ESTIMACIÓN POLIMETÁLICA EMPLEANDO TÉCNICAS DE INTELIGENCIA COMPUTACIONAL



PAMELA ANTONIA MELLA FIGUEROA

Informe de Memoria de Título

para optar al Título de

Ingeniero Civil de Minas

AGOSTO 2020



A mis padres, Francisco y Pamela

AGRADECIMIENTOS

En primer lugar, quiero agradecer a mi familia, especialmente a mis padres, que estuvieron presentes en todo mi proceso universitario, ya sea dándome su apoyo en los momentos difíciles, llevándome tecitos en las noches mientras estudiaba, prendiendo la estufa para que no pasara frio las veces que me amanecía estudiando, llevándome a la universidad en auto los días de lluvia para que no me mojara, y así un sinfín de cosas que definitivamente me alivianaron la carga. A mis hermanos, por darme ánimo las veces en que la carrera se ponía más difícil, gracias por siempre alegrarme la vida.

A mis amigos, los que conocí en mi primer año de Ingeniería Civil Plan Común, que siendo el primer año siempre el más difícil, lograron que fuera un año lleno de risas y buenos momentos que recuerdo siempre. A los que luego conocí en Civil de Minas, Pauli, Vale y Cris, gracias por todas esas risas, papitas del Sabrossón y noches de estudios en la central, a pesar de que tantas veces nos vimos colapsados con un ramo, “las risas no faltaron”.

A mi Chikibeibi mailob, Rafael, le agradezco a la vida esa pasantía a Australia que nos unió, y a pesar de que no la hubiésemos hecho juntos, de una u otra forma la vida se hubiese encargado de juntarnos. Gracias por todo tu apoyo, por tu contención, por sacarme risas en mis momentos de colapso, por las papitas de la Cocina, por ayudarme prestándome todos los computadores que había en tu casa para alivianar mi trabajo, sin ti quizás cuanto me hubiese demorado en terminar la memoria, gracias mi amor.

A mis profesores, a mi profesor guía y a todos aquellos que de alguna u otra manera me ayudaron con una palabra de aliento, muchas gracias.

RESUMEN

La estimación de recursos minerales es esencial para las etapas posteriores de un proyecto minero, tales como el diseño y planificación, pues una mala ejecución de esta conlleva a pérdidas monetarias significativas. Técnicas geoestadísticas tradicionales de estimación, como Kriging y sus variantes, son las populares para ser usadas en la industria minera, sin embargo, cuando el yacimiento presenta un comportamiento geológico complejo, los resultados obtenidos de la estimación geoestadística no son los mejores. Los recientes avances que han presentado las técnicas de Inteligencia Computacional, han hecho de este un foco de atención, lo que ha provocado una serie de investigaciones para la estimación de leyes minerales. Es debido a esto que se propone utilizar los algoritmos Redes Neuronales Artificiales, Árboles de Regresión y *Support Vector Regression* para ser comparados con el método tradicional Cokriging Ordinario en situaciones donde se tenga baja correlación espacial, poca información de los datos y una alta correlación entre variables.

La metodología propuesta consta en dos casos de estudio: uno simulado y uno real. El caso simulado consiste en realizar cien simulaciones gaussianas de tres variables continuas pertenecientes a un yacimiento, donde cada base generada es particionada en una base para entrenar los algoritmos y otra para realizar la validación de estos. El segundo caso corresponde a un yacimiento real, que contiene dos variables continuas, oro y cobre, además de cuatro dominios geometalúrgicos, donde estos se extienden en un modelo de bloques para realizar la estimación de manera independiente. Una vez obtenido los resultados en los dos casos, se realiza la comparación de los métodos en cada caso individualmente y luego entre uno y otro. Para ambos casos se realiza un plan de categorización y se mide la calidad de esta, finalizando con una evaluación económica.

Los resultados obtenidos, tanto en el caso simulado como en el real, indican que el método *Support Vector Regression* obtiene mejores resultados en cuanto a la disminución del error de estimación en comparación con el método tradicional. Para los casos simulados, el 95% de las veces *Support Vector Regression* obtiene mejores resultados que Cokriging Ordinario, disminuyendo el error de estimación en aproximadamente un 7%, mientras que, para el caso real, en el dominio en que se replican las condiciones señaladas en un inicio, se logra una disminución del 8%. Además, este método junto a Árboles de Regresión, parecen ser los más conservadores a la hora de estimar.

ABSTRACT

Mineral resource estimation is essential in the later mine project steps, such as planification and mine design, because bad execution of those imply significant income loses. Traditional geostatistical estimation techniques, like Kriging and its variants, are amongst the most popular used on the mining industry, nevertheless, when the deposit presents a complex geological behavior, the results obtained by these methods aren't the best. Recent developments obtained by Computational Intelligence techniques, have made themselves a main attraction, which have produced a series of investigations related to ore estimation using the aforementioned. Due to this, it is proposed to use algorithms from this school: Artificial Neural Networks, Regression Trees and Support Vector Regression, to compare them with traditional Ordinary Cokriging in situations when low spatial correlation, little information is available and high correlation between variables is present.

The proposed methodology implies the study of two cases: simulated and real cases. The simulated case consists in making a hundred gaussian simulations, with three continuous variables each, forming part of a deposit where every database is split between training and validation to test these algorithms. The second case corresponds to a real deposit, which contains two continuous variables, gold and copper, in addition to four geometallurgical domains, whose categories are extended to the block model to realize estimation independently. Once results are produced in both cases, a comparison is done in each one individually and between them. A categorization plan, with its respective quality, is carried out for both cases, finishing with an economic evaluation.

The results obtained, both in the simulated case as in the real one, suggest that Support Vector Regression manages to produce better results in decreasing the error obtained when comparing the Computational Intelligence algorithms with the traditional ones. In the simulated cases, 95% of the times Support Vector Regression produces better results than Ordinary Cokriging, lowering the estimation error by a 7%, meanwhile, in the real case, in the third domain, where hypothesis conditions exist, a decrease of 8% in the error is achieved. Besides, this method, along with Regression Trees, seem to be the most conservatives at the time of estimating.

TABLA DE CONTENIDOS

CAPÍTULO I: INTRODUCCIÓN.....	1
1.1 Introducción	1
1.2 Descripción del problema.....	3
1.3 Hipótesis de trabajo	4
1.4 Objetivos.....	4
1.4.1 Objetivos generales	4
1.4.2 Objetivos específicos.....	5
1.5 Metodología	5
1.6 Alcances	6
CAPÍTULO II: MARCO TEÓRICO	7
2.1 Cokriging ordinario	7
2.1.1 Definiciones básicas de la geoestadística	8
2.1.2 Diseño de la vecindad de cokriging	13
2.1.3 Hipótesis del cokriging ordinario.....	14
2.1.4 Principios del cokriging ordinario.....	14
2.1.5 Determinación del estimador	16
2.1.6 Varianza del cokriging ordinario	16
2.2 Redes Neuronales	17
2.2.1 Noción de neurona biológica	17
2.2.2 Neurona artificial	17
2.2.3 Red neuronal artificial.....	19
2.2.4 Estructura de una ANN	20
2.2.5 Aprendizaje en las ANN.....	21
2.2.6 Algoritmo <i>backpropagation</i>	24
2.3 Árboles de Regresión.....	25
2.3.1 Método CART	26
2.3.2 Elementos de un árbol	27
2.3.1 Particionamiento recursivo	27
2.3.2 Impureza del nodo.....	28
2.3.3 Reglas de división de un nodo	28

2.3.4	Podado del árbol (<i>pruning</i>).....	29
2.3.5	Costo-complejidad	30
2.4	<i>Support Vector Machine</i>	30
2.4.1	SVR caso lineal.....	31
2.4.2	SVR caso no lineal.....	35
2.4.3	Funciones de costo	38
2.5	Categorización de recursos minerales	41
2.5.1	Metodología de categorización de recursos minerales	42
CAPÍTULO III: CASO SIMULADO		44
3.1	Metodología de trabajo	44
3.2	Aplicación	45
3.2.1	Generalización del espacio de trabajo	45
3.2.2	Simulación de variables continuas	46
3.2.3	Particionamiento de base.....	49
3.2.4	Métodos de estimación	50
3.2.5	Medidas de error de estimación	62
3.2.6	Categorización de recursos minerales	65
3.2.7	Análisis económico	69
3.3	Análisis de resultados	73
CAPÍTULO IV: CASO REAL.....		79
4.1	Metodología de trabajo	79
4.2	Estudio exploratorio de datos	80
4.2.1	Datos duplicados y <i>outliers</i>	82
4.3	Aplicación	84
4.3.1	Creación del modelo de bloques	84
4.3.2	Particionamiento de base.....	86
4.3.3	Métodos de estimación	87
4.3.4	Medidas de error de estimación	95
4.3.5	Medidas estadísticas principales del modelo de bloques	97
4.3.6	Categorización de recursos minerales	99
4.3.7	Análisis económico	101
4.4	Análisis de resultados	104
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES.....		108

REFERENCIAS	111
ANEXO A: SECCIÓN DE REDES NEURONALES ARTIFICIALES	114
Anexo A. 1: Funciones de activación para las ANN.....	114
Anexo A. 2: Tipos de ANN.....	117
ANEXO B: SECCIÓN DE SVM.....	119
Anexo B. 1: Introducción a la Teoría de Optimización	119
Anexo B. 2: Transformación a problema dual.....	120
Anexo B. 3. Cálculo de b.....	121
ANEXO C: LEY EQUIVALENTE Y LEY DE CORTE	123
Anexo C. 1: Ley equivalente.....	123
Anexo C. 2: Ley de corte	123
ANEXO D: CASO SIMULADO	125
Anexo D. 1: <i>Skill value</i>	125
Anexo D. 2: Evaluación económica	128
ANEXO E: CASO REAL.....	130
Anexo E. 1: Evaluación económica	130

ÍNDICE DE FIGURAS

Figura 1. Etapas de un proyecto minero.	1
Figura 2. Modelo de neurona artificial.	18
Figura 3. Modelo de red neuronal artificial con las conexiones hacia delante.	19
Figura 4. Ejemplo de red <i>feedforward</i>	20
Figura 5. Ejemplos de red <i>feedback</i>	21
Figura 6. Modelo de árbol de regresión.	27
Figura 7. SVR con margen blando.	33
Figura 8. Funciones de pérdida.	40
Figura 9. Relación entre recursos minerales y reservas de mena.	42
Figura 10. Espacio de trabajo.	45
Figura 11. Variogramas simples y cruzados utilizados para hacer la simulación de las variables continuas.	47
Figura 12. Leyes de cobre, molibdeno y oro de la simulación uno.	48
Figura 13. Sondajes extraídos de la simulación uno.	49
Figura 14. Gráfico de densidad para la base total, de entrenamiento y validación de la simulación uno.	50
Figura 15. Mapas variográficos de la base de entrenamiento de la simulación uno.	51
Figura 16. Variogramas omnidireccionales de la base de entrenamiento para la simulación uno.	52
Figura 17. Estimación con Cokriging Ordinario para la simulación uno.	53
Figura 18. Configuración de ANN para la simulación uno.	56
Figura 19. Estimación con Redes Neuronales Artificiales para la simulación uno.	57
Figura 20. Árbol de decisión para la simulación uno.	59
Figura 21. Estimación con Árboles de Regresión para la simulación uno.	60
Figura 22. Estimación con <i>Support Vector Regression</i> para la simulación uno.	62
Figura 23. Gráfico de caja de los <i>skill value</i> obtenidos para cada método en las cien simulaciones.	64
Figura 24. Categorización de recursos minerales para la estimación del mejor, peor y más probable caso.	66

Figura 25. <i>Boxplot</i> de las leyes de cobre y oro para la base real en general y según dominios.....	81
Figura 26. Visualización de leyes de cobre y oro, además de los dominios presentes en la base real.	82
Figura 27. <i>Boxplot</i> de las leyes de cobre y oro para la base real en general y según dominios después de limpiarla.	84
Figura 28. Modelo de bloques para la base real.	85
Figura 29. Extensión de dominios geometalúrgicos en modelo de bloques.	86
Figura 30. Distribución de la base de datos real, de entrenamiento y validación.	87
Figura 31. Mapas variográficos de la base real según dominio.	88
Figura 32. Variogramas direccionales de la base real según dominio.	89
Figura 33. Estimación del oro con Cokriging Ordinario en la base real, dominio 4, cota 299.4.....	90
Figura 34. Estimación del oro con Redes Neuronales Artificiales en la base real, dominio 4, cota 299.4.	92
Figura 35. Estimación del oro con Árboles de Regresión en la base real, dominio 4, cota 299.4.	93
Figura 36. Estimación del oro con <i>Support Vector Regression</i> en la base real, dominio 4, cota 299.4.	95
Figura 37. Categorización de recursos minerales para la estimación de la base real de validación y modelo de bloques.	100
Figura 38. Estimación de dominio 2 en el modelo de bloques a diferentes cotas.....	107

ÍNDICE DE TABLAS

Tabla 1. Ejemplos de funciones kernels.	38
Tabla 2. Funciones de pérdida comunes y sus correspondientes modelos de densidad.	41
Tabla 3. Clasificación de recursos minerales.	43
Tabla 4. Parámetros de variogramas simples y cruzados utilizados para crear las simulaciones de variables continuas.	46
Tabla 5. Correlación entre el cobre, molibdeno y oro en la base de entrenamiento para la simulación uno.	55
Tabla 6. Parámetros óptimos de ANN para la simulación uno.	55
Tabla 7. Parámetros óptimos de CART para la simulación uno.	59
Tabla 8. Parámetros óptimos de SVR para la simulación uno.	61
Tabla 9. Medidas de rendimiento en las estimaciones para el mejor, peor y más probable caso.	65
Tabla 10. Calidad de los datos para la estimación del mejor, peor y más probable caso según recursos minerales.	68
Tabla 11. Ranking de la calidad de datos para estimación del mejor, peor y más probable caso según recursos minerales.	69
Tabla 12. Datos económicos considerados para los casos simulados.	69
Tabla 13. Ley media y tonelaje utilizando ley equivalente para estimación del mejor, peor y más probable caso.	70
Tabla 14. Ley media y tonelaje después de filtrar con ley de corte para estimación del mejor, peor y más probable caso.	71
Tabla 15. Evaluación económica para estimación del mejor, peor y más probable caso.	72
Tabla 16. Diferencia entre los beneficios de la base original y la base estimada para el mejor, peor y más probable caso.	72
Tabla 17. Porcentaje de veces en las que el método de estimación utilizado obtiene mejor y peor <i>skill value</i>	74
Tabla 18. Porcentaje de veces en las que ML obtiene mejor <i>skill value</i> que Cokriging Ordinario. ...	74
Tabla 19. Promedio de porcentaje de mejoramiento y empeoramiento de ML versus Cokriging Ordinario.	75

Tabla 20. Promedio de la calidad de los datos estimados en las cien simulaciones según categoría de recursos.	76
Tabla 21. Porcentaje de veces en la que un método es más o menos conservador según categoría de recursos.	76
Tabla 22. Porcentaje de veces en las que un método es más o menos conservador en general.	77
Tabla 23. Porcentaje de veces en las que un método se asemeja o aleja del beneficio económico real.	78
Tabla 24. Porcentaje de veces en que ML es mejor que OCOK en cuanto al error producido en el beneficio económico.	78
Tabla 25. Coordenadas máximas y mínimas de la base de datos real.	80
Tabla 26. Medidas estadísticas principales de las leyes de cobre y oro para cada dominio de la base real.	80
Tabla 27. Medidas estadísticas principales de la base real según dominio después de limpiarla.	83
Tabla 28. Coordenadas máximas y mínimas del modelo de bloques.	84
Tabla 29. Cantidad de aciertos en la predicción de dominios para la base real.	86
Tabla 30. Parámetros de los modelos ajustados de variogramas simples y cruzados para la base real según dominio.	90
Tabla 31. Parámetros óptimos de ANN para la base real según dominio.	91
Tabla 32. Parámetros óptimos de CART para la base real según dominio.	93
Tabla 33. Parámetros óptimos de SVR para el caso real según dominio.	94
Tabla 34. Medidas de rendimiento en las estimaciones para el caso real según dominio.	96
Tabla 35. Medidas de rendimiento en las estimaciones para el caso real de manera global.	97
Tabla 36. Estadísticas principales de las leyes estimadas en el modelo de bloques según dominio.	98
Tabla 37. Estadísticas principales de las leyes estimadas en el modelo de bloques de manera global.	98
Tabla 38. Calidad de los datos para la estimación de la base real según recursos minerales.	101
Tabla 39. Ranking de la calidad de los datos para la estimación de la base real según recursos minerales.	101
Tabla 40. Datos económicos considerados para el caso real.	102
Tabla 41. Ley media y tonelaje utilizando ley equivalente para estimación de bloques de manera global.	102

Tabla 42. Ley media y tonelaje después de filtrar con ley de corte para estimación de bloques de manera global.	103
Tabla 43. Evaluación económica para estimación de bloques de manera global.	103
Tabla 44. Mejor <i>skill value</i> para la base real según dominio.	104
Tabla 45. Porcentaje de mejoramiento y empeoramiento de ML versus Cokriging Ordinario.	105
Tabla 46. Ranking de la calidad de los datos para la estimación de la base real en general.	105
Tabla 47. Principales funciones de activación.	114
Tabla 48. Tipos de Redes Neuronales Artificiales más importantes.	117
Tabla 49. <i>Skill values</i> para cobre, molibdeno y oro en las cien simulaciones.	125
Tabla 50. Diferencia entre los beneficios de la base original y la base estimada en las cien simulaciones.	128
Tabla 51. Evaluación económica para estimación de bloques según dominio.	130



NOMENCLATURA

IC	Inteligencia Computacional
IA	Inteligencia Artificial
ML	Aprendizaje automático, <i>Machine Learning</i>
MLC	Modelo lineal de correogionalización
ANN	Red Neuronal Artificial, <i>Artificial Neural Networks</i>
CART	Árboles de clasificación y regresión, <i>Classification and Regression Trees</i>
SVM	Máquina de soporte vectorial, <i>Support Vector Machine</i>
SVR	<i>Support Vector Regression</i>
ANOVA	Análisis de varianza, <i>Analysis of Variance</i>
ME	Error medio, <i>Mean Error</i>
MAE	Error absoluto medio, <i>Mean Absolute Error</i>
R^2	Coficiente de determinación
RMSE	Raíz del error cuadrático medio, <i>Root Mean Square Error</i>
COG	Ley de corte, <i>cut off grade</i>



CAPÍTULO I: INTRODUCCIÓN

1.1 INTRODUCCIÓN

Los depósitos minerales son una concentración anómala de un determinado mineral con respecto a su entorno. Que este pase a ser un yacimiento para operar como una mina quiere decir que su explotación producirá un beneficio económico en un determinado momento. Para que esto ocurra, es necesario el desarrollo de un proyecto minero, el cual consta de una serie de extensas etapas, las que pueden ser contempladas en cinco fases: prospección, exploración, desarrollo, explotación y recuperación [1] (Figura 1). La primera etapa consiste en la búsqueda de minerales valiosos, donde se utilizan técnicas directas e indirectas, las cuales están constituidas por un examen visual de la superficie y por el empleo de la herramienta científica de la geofísica. Una vez terminada la prospección, se desarrolla la exploración con la finalidad de determinar tanto el tamaño y contenido del depósito mineral, como su valor, para así realizar una estimación sobre costos, recuperación, entre otros, a fin de determinar la rentabilidad de este. La siguiente etapa de desarrollo se realiza con el objetivo de acceder al mineral para su posterior explotación. Una vez que se tenga el suficiente desarrollo como para comenzar la producción sin interrupciones, comienza la explotación de la mena. La última etapa, llamada recuperación, es el proceso de cierre de la mina, contemplando la revegetación y restauración de la tierra y agua.



Figura 1. Etapas de un proyecto minero.

Con la información recopilada en las fases de prospección y exploración, se realiza la estimación del recurso mineral¹. Una estimación en minería consiste en la determinación de las características de

¹ Los recursos minerales son definidos como una concentración de material en o sobre la corteza terrestre de tal manera que su forma y cantidad tengan una expectativa razonable para su extracción económica. Su ubicación, cantidad, ley, características geológicas y continuidad son conocidas, estimadas o interpretadas con la ayuda de evidencias y conocimiento geológico [2].

un mineral del que no se tiene conocimiento a partir de sondajes de perforación disponibles, de esa manera resulta ser una de las plataformas de mayor importancia sobre las cuales se desarrolla un proyecto minero, ya que su influencia es crítica para que las siguientes etapas tengan una buena ejecución, por lo que se busca que sea lo más exacta y precisa posible.

Existen métodos tradicionales que estiman leyes de un depósito mineral estudiando la información de una sola variable, tales como el vecino más cercano, inverso de la distancia, *splines* y kriging [3], donde los primeros tres son métodos determinísticos y este último pertenecen a la ciencia de la geoestadística², lo que puede considerarse como una limitación pues es debido considerar que los depósitos minerales son originados por diversos procesos genéticos de formación, definiendo su caracterización mineralógica, geoquímica y estructural, donde este puede favorecer la ocurrencia de elementos con un comportamiento químico similar, haciendo que estén asociados y se correlacionen entre sí [4]. Es así como estos elementos aportan información y de esta manera pueden ser útiles en la estimación de una variable. Por ejemplo, un depósito mineral que contenga sulfuros de cobre contiene también compuestos de otros metales tales como el hierro, zinc, arsénico, entre otros.

Dado que la precisión de las leyes es clave para el diseño y desarrollo de una planificación minera, es importante utilizar la información adicional disponible en caso de que sea relevante. El cokriging es una extensión del kriging que incorpora la información de varias variables para realizar las estimaciones. El rendimiento de este es mucho mejor cuando la correlación espacial es buena y la relación existente entre los patrones de entrada y salida es lineal. Si bien existen variantes del kriging que se adaptan a una relación no lineal, no logran capturar eficientemente la naturaleza de la no linealidad espacial.

Por otro lado, existen técnicas de Inteligencia Computacional (IC), rama de la Inteligencia Artificial (IA), que se utilizan para realizar estimaciones con relaciones no lineales, sin tener la necesidad de que la correlación espacial sea alta para lograr resultados beneficiosos. En la actualidad, la IC vive un creciente interés pues todas las industrias quieren tomar decisiones mejor informadas y más efectivas, además de reducción de costos y ahorro de tiempo. En este contexto, existe un

² La geoestadística es una rama de la estadística aplicada que se encarga de estudiar fenómenos regionalizados. Utiliza un conjunto de técnicas de predicción y simulación para ser desarrolladas en la disciplina de las ciencias de la tierra [5].

subcampo que se denomina *Machine Learning* (ML) o aprendizaje automático que se dirige al estudio y desarrollo de algoritmos con la capacidad de aprendizaje a partir de datos u observaciones, donde la tarea principal es hacer inferencia a partir de una muestra [6]. En el transcurso de los años se han desarrollado distintas metodologías en las que se pueden predecir valores, ya sea por clasificación o estimación [7]. Algunas de estas técnicas son: Redes Neuronales, Algoritmos Genéticos, Árboles de Decisión, Máquinas de Soporte Vectorial, Regresión Lineal, entre otros.

La principal diferencia entre las técnicas de geoestadística y de IC, es que las primeras favorecen la interpretabilidad del modelo estadístico, mientras que las últimas enfatizan en el rendimiento de la predicción y el rendimiento computacional de los algoritmos [6].

Existen estudios en los que se realizan comparaciones entre estimaciones utilizando algoritmos de ML y métodos geoestadísticos, donde en muchos de los casos existen mejores resultados para las primeras [8-11]. Sin embargo, existen pocos estudios para el caso multivariable, en los que generalmente se centran en casos particulares, y no existen investigaciones que consideren la categorización de recursos y posterior análisis económico para representar el impacto monetario, que finalmente es lo relevante en un proyecto minero.

Este trabajo se fundamenta en la estimación de leyes de un depósito mineral polimetálico considerando una baja correlación espacial y alta correlación entre los minerales. La base es emplear principalmente el método geoestadístico Cokriging y los algoritmos de ML llamados Máquinas de Soporte Vectorial, Árboles de Regresión y Redes Neuronales Artificiales. La estructura de la investigación consiste en programar las técnicas mencionadas en el software RStudio, las que posteriormente son comparadas, evaluando la categorización de recursos para cada método utilizado y finalizando en un análisis económico.

1.2 DESCRIPCIÓN DEL PROBLEMA

El problema abordado en esta memoria de título está relacionado con la estimación de recursos minerales de un yacimiento polimetálico y posterior categorización utilizando herramientas de Inteligencia Computacional.

En ocasiones, existen casos desfavorables donde el comportamiento de las variables metalúrgicas es errático, mostrando poca o nula correlación espacial. En estos casos, y adicionalmente considerando la usual escasa información proveniente de las campañas de exploración, la cuantificación de recursos minerales se ve ampliamente afectada, pues la técnica popular de geoestadística que se utiliza para hacer la estimación de recursos precisa de la correlación espacial de los datos, además de la cantidad de estos, provocando que el error a posteriori este directamente relacionado con estos factores.

Bajo este contexto, el cokriging parece ser desfavorable para ser utilizado como método de estimación, ya que, si bien este utiliza información adicional a la variable primaria, al existir baja correlación espacial no aprovecha toda la información valiosa que se pueda tener en las bases de datos.

Una mala precisión en la estimación conduce a un alto riesgo para las etapas siguientes de planificación y diseño minero, pues todas las decisiones tomadas son basadas en los resultados de esta, implicando pérdidas monetarias significativas en el proyecto minero.

1.3 HIPÓTESIS DE TRABAJO

Los algoritmos estadísticos basados en la Inteligencia Computacional resuelven el problema de la estimación de recursos minerales ante una baja o nula correlación espacial y poca información de las variables proveniente de los sondeos, considerando la existencia de una buena correlación entre las variables a estimar. De esta forma, es posible realizar una categorización de recursos más confiable y generar menos riesgo económico en un proyecto minero.

1.4 OBJETIVOS

1.4.1 OBJETIVOS GENERALES

El objetivo general de esta memoria de título es aplicar diferentes algoritmos, incorporando modelos de geoestadística y de inteligencia computacional, para estimar las leyes de un depósito

mineral polimetálico y evaluar cómo estos impactan en la evaluación del riesgo económico de un proyecto minero.

1.4.2 OBJETIVOS ESPECÍFICOS

- Aplicar algoritmos de Inteligencia Computacional para estimar variables continuas con pocos datos y baja correlación espacial.
- Crear modelo de las leyes de los diferentes minerales presentes en el cuerpo mineralizado.
- Evaluar categorización de recursos según el método de estimación utilizado.
- Realizar un análisis económico de diferentes escenarios de riesgo.
- Comparar las técnicas de estimación multivariable utilizadas en el estudio: cokriging ordinario, máquinas de soporte vectorial, árboles de regresión y redes neuronales.

1.5 METODOLOGÍA

La metodología planteada para abordar el problema descrito consiste en el análisis de dos casos de estudio: una base con datos simulados y otra base con datos reales.

El procedimiento que se realiza para trabajar con el conjunto de datos simulados se contempla en los siguientes pasos:

- Generar una base de datos simulando tres variables continuas que representan al cobre, molibdeno y oro, las que van a estar distribuidas en un espacio 3D, correspondiente a un muestreo de una zona de estudio.
- Particionar la base simulada en dos, donde una parte corresponde a los datos de entrenamiento y la otra a los datos de validación. Con la primera se realizan predicciones de las variables regionalizadas continuas previamente mencionadas en las localidades de los datos de la base de datos correspondiente a la validación, donde se realiza una comparación entre los valores predichos y los valores reales.
- Estimar las variables continuas utilizando Cokriging Ordinario y métodos de Inteligencia Computacional.

- Realizar un estudio comparativo con las medidas estadísticas principales asociadas a las estimaciones realizadas, estudiando la variabilidad de las predicciones.
- Para cada predicción, generar un plan de categorización de recursos minerales, reportando las medidas estadísticas principales.
- Realizar análisis económico asociados a los resultados obtenidos y analizar cómo este se diferencia del escenario original.

Por otro lado, el procedimiento a realizar para el desarrollo del trabajo con el conjunto de datos de una base real es el siguiente:

- Realizar una descripción detallada de la base de datos.
- Realizar un análisis descriptivo de cada una de las variables presentes en la base de datos.
- Estimar las variables utilizando algoritmos de métodos tradicionales y de Inteligencia Computacional.
- Para las predicciones resultantes, hacer un plan de categorización de recursos indicando las medidas estadísticas principales.
- Realizar evaluación económica con cada resultado obtenido.

Para ambos casos, se finaliza con un análisis crítico de los resultados, elaborando conclusiones y discusiones parciales.

1.6 ALCANCES

La investigación es desarrollada utilizando dos bases de datos para realizar la estimación polimetálica, un caso sintético (base simulada) y un caso real, donde en ambos se trabaja con pocos sondeos de una localidad con baja correlación espacial, cuyos resultados obtenidos son en base a estas condiciones específicas. La estimación consiste en aplicar los algoritmos Cokriging Ordinario, Redes Neuronales Artificiales, Árboles de Regresión y Máquinas de Soporte Vectorial en el software RStudio según la capacidad computacional disponible y, una vez obtenidos los resultados, la comparación se realiza entre dichos métodos.

CAPÍTULO II: MARCO TEÓRICO

El siguiente capítulo busca describir los mecanismos que son utilizados en el desarrollo de la investigación, enfocado principalmente en explicar la técnica geoestadística de cokriging y las técnicas de IC compuestas por Redes Neuronales Artificiales, Árboles de Regresión y Máquinas de Soporte Vectorial.

2.1 COKRIGING ORDINARIO

El kriging es un método de estimación espacial univariable, integrado por primera vez por Danie G. Krige [12] y formalizados en la década de los 60 por Georges Matheron [13], cuyas bases se fundamentan en la geoestadística. Esta técnica nace con la intención de integrar información que los métodos tradicionales de interpolación no consideran, como la continuidad espacial de la variable de estudio, y su principal fundamento es minimizar el error cuadrático medio de la estimación [14].

Las diferentes variables del kriging se articulan en torno a las mismas etapas, donde su diferencia radica en la hipótesis que se hace sobre la función aleatoria. El cokriging es una extensión del kriging para el caso de estimación multivariable, tomando en cuenta la continuidad espacial de estas y sus relaciones de dependencia entre ellas, esto es, además de considerar la variable de interés, considera la información de variables adicionales, mejorando su precisión. Para realizar la estimación multivariable, es necesario considerar tanto los variogramas (o covarianzas) individuales de cada variable (variogramas/covarianzas simples) como los variogramas (o covarianzas) cruzados entre ellas. El modelo es ventajoso cuando la variable de interés (variable primaria) está sub-mostrada con respecto a las otras variables que tienen correlación con ella (covariables o variables secundarias) [15].

Para tener una mejor comprensión de la técnica de cokriging, primero es necesario definir algunos conceptos básicos de la geoestadística.

2.1.1 DEFINICIONES BÁSICAS DE LA GEOESTADÍSTICA

La geoestadística es una rama de la estadística que trata fenómenos espaciales, la cual ofrece maneras de describir la continuidad espacial, adaptando las técnicas tradicionales a esta. Sus principales objetivos son la predicción, estimación y simulación de estos fenómenos espaciales [14], los que pueden ser aplicados en la evaluación de recursos naturales, topografía, agricultura, etc.

Los fundamentos básicos utilizados en la etapa de estimación por cokriging ordinario son presentados a continuación.

2.1.1.1 Variable regionalizada

Una variable regionalizada es una variable medida que se extiende en el espacio y presenta una estructura de correlación, la cual representa propiedades o atributos asociados a un fenómeno natural, como, por ejemplo, la ley de un mineral, la concentración de un contaminante, el tipo de roca, etc. Esta se denota como $z(x)$ donde x representa la posición en el espacio y dado que esta no posee una extensión infinita, se estudia dentro de un dominio limitado llamado campo de la variable D [5].

En general, la variable regionalizada no es conocida en su totalidad, sino que se cuenta con un conjunto de datos que la describe (muestras, observaciones, mediciones, etc.). Es en este contexto donde se aprovecha la información disponible y se construyen modelos que permitan describir y entender a la variable regionalizada.

2.1.1.2 Función aleatoria

El valor $z(x)$ de la variable regionalizada en el sitio x del campo D se interpreta como una realización de una variable aleatoria $Z(x)$. El conjunto de variables aleatorias constituye una función aleatoria que está dada por $Z = \{Z(x), x \in D \subset \mathbb{R}^d\}$. Luego, para cualquier k argumentos y una serie de sitios x_1, x_2, \dots, x_k el vector aleatorio $\vec{Z}(x) = [Z(x_1), Z(x_2), \dots, Z(x_k)]^T$, formado por un grupo de variable aleatorias, está definido por una distribución espacial, la que está formada por el conjunto de funciones de distribución como sigue [14]:

$$F_{x_1, x_2, \dots, x_k}[z_1, z_2, \dots, z_k] = P[Z(x_1) \leq z_1, Z(x_2) \leq z_2, \dots, Z(x_k) \leq z_k] \forall z_1, z_2, \dots, z_k \in \mathbb{R} \quad (1)$$

Para determinar la distribución espacial de las funciones aleatorias, son necesarias algunas restricciones:

1. **Estacionaridad:** postula que la distribución espacial de las funciones aleatorias es invariante respecto a cualquier traslación de un vector h . De esa manera, las propiedades de un conjunto de datos no dependen su posición absoluta en el espacio [5].
2. **Estacionaridad de segundo orden:** se dice que una variable aleatoria es estacionaria de segundo orden si su valor esperado es finito y constante para todo D y si, además, la covarianza existe para un vector de traslación h [14].

$$E[Z_i(x)] = m_i \quad (2)$$

$$\text{cov}[Z_i(x+h), Z_j(x)] = C_{ij}(h) \quad (3)$$

3. **Estacionaridad intrínseca:** cuando la varianza no es finita, se trabaja con la hipótesis de que los incrementos sean estacionarios. Esto ocurre si la esperanza de los incrementos definidos por un vector h es cero y la varianza del incremento está definida y es función única de este [14].

$$E[Z_i(x+h) - Z_i(x)] = 0 \quad (4)$$

$$\text{var}[Z_i(x+h) - Z_j(x)] = 2\gamma_{ij}(h) \quad (5)$$

2.1.1.3 Covarianza simple y cruzada

La covarianza es un indicador que muestra la dependencia de una variable en el espacio, es decir, señala qué tan semejantes son los valores medidos. La covarianza simple se refiere al caso univariable, mientras que la cruzada es para el caso en el que existen covariables, donde se representa la relación de dependencia entre ellas.

La covarianza simple está definida como sigue:

$$C(h) = \text{cov}[Z(x+h), Z(x)] = E\{[Z(x+h) - m][Z(x) - m]\}, \quad (6)$$

donde m es el valor esperado. Con una variable regionalizada z conocida en n sitios $\{x_1, x_2, \dots, x_n\}$, un estimador de la covarianza simple se obtiene al reemplazar m por la media aritmética sobre los pares de datos, donde la separación de estos es un vector h :

$$\hat{C}(h) = \frac{1}{|N(h)|} \sum_{N(h)} [z(x_\alpha) - \bar{z}][z(x_\beta) - \bar{z}], \quad (7)$$

donde $N(h) = \{(\alpha, \beta) \text{ tal que } x_\alpha - x_\beta = h\}$ y $|N(h)|$ es el número de pares contenidos en el conjunto $N(h)$. Por otro lado, \bar{z} es la media aritmética de los datos disponibles.

La covarianza cruzada entre dos funciones aleatorias Z_i y Z_j para un vector h se define según la ecuación (8), mientras que su estimador se expresa según la ecuación (9).

$$C_{ij}(h) = \text{cov}[Z_i(x+h), Z_j(x)] = E\{Z_i(x+h)Z_j(x)\} - E\{Z_i(x+h)\} \times E\{Z_j(x)\} \quad (8)$$

$$\hat{C}_{ij}(h) = \frac{1}{|N_{ij}(h)|} \sum_{N_{ij}(h)} [z_i(x_\alpha^i) - \bar{z}_i][z_j(x_\beta^j) - \bar{z}_j], \quad (9)$$

con

$$\bar{z}_i = \frac{1}{n_i} \sum_{\alpha=1}^{n_i} z_i(x_\alpha^i) \quad \text{y} \quad \bar{z}_j = \frac{1}{n_j} \sum_{\beta=1}^{n_j} z_j(x_\beta^j), \quad (10)$$

donde $N_{ij}(h) = \{(\alpha, \beta) \text{ tal que } x_\alpha^i - x_\beta^j = h\}$ y $|N_{ij}(h)|$ es el número de pares distintos del conjunto $N_{ij}(h)$ y x_α^i y x_β^j son las posiciones donde fueron medidas las variables Z_i y Z_j , respectivamente.

A continuación, se presentan algunas de las propiedades de la covarianza cruzada.

1. $C_{ij}(h) = C_{ji}(-h)$ (11)

2. $C_{ii}(0) = \sigma_i^2$ (12)

$$3. C_{ii}(0)C_{jj}(0) \geq |C_{ij}(h)|^2, \quad (13)$$

donde σ_i^2 corresponde a la varianza de Z_i . La ventaja de la covarianza cruzada sobre el variograma cruzado, es que la primera puede ser utilizada independientemente de la configuración de muestreo [16].

2.1.1.4 Variograma simple y cruzado

El objetivo del variograma es cuantificar la variabilidad de una medición en el espacio, indicando, en simples palabras, que tan desemejantes son los datos medidos. Al igual que para la covarianza, el simple es para el caso univariable y el cruzado es para el caso multivariable, representando la relación espacial entre dos variables.

El variograma teórico simple de una función aleatoria Z con un vector de separación h se presenta a continuación.

$$\gamma(h) = \frac{1}{2} E\{[Z(x+h) - Z(x)]^2\} \quad (14)$$

Con una variable regionalizada z conocida en n sitios $\{x_1, x_2, \dots, x_n\}$, el estimador del variograma simple para un vector h es:

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} [z(x_\alpha) - z(x_\beta)]^2, \quad (15)$$

donde $N(h) = \{(\alpha, \beta) \text{ tal que } x_\alpha - x_\beta = h\}$ y $|N(h)|$ es el número de pares contenidos en el conjunto $N(h)$.

El variograma cruzado entre las funciones aleatorias Z_i y Z_j para un vector h se define según la ecuación (16) y su estimador queda definido por la ecuación (17).

$$\begin{aligned}\gamma_{ij}(h) &= \frac{1}{2} \text{cov}[Z_i(x+h) - Z_i(x), Z_j(x+h) - Z_j(x)] \\ &= \frac{1}{2} E\{[Z_i(x+h) - Z_i(x)][Z_j(x+h) - Z_j(x)]\}\end{aligned}\tag{16}$$

$$\hat{\gamma}_{ij}(h) = \frac{1}{2|N_{ij}(h)|} \sum_{N_{ij}(h)} [z_i(x_\alpha) - z_i(x_\beta)][z_j(x_\alpha) - z_j(x_\beta)],\tag{17}$$

donde $N_{ij}(h) = \{(\alpha, \beta) \text{ tal que } x_\alpha^i - x_\beta^j = h, \text{ siendo a la vez } z_i \text{ y } z_j \text{ medidas en } x_\alpha \text{ y } x_\beta\}$.

Algunas de las propiedades del variograma cruzado son:

1. $\gamma_{ij}(0) = 0$ (18)

2. $\gamma_{ij}(h) = \gamma_{ij}(-h)$ (19)

3. $\gamma_{ij}(h) = \gamma_{ji}(h)$ (20)

4. $\gamma_{ii}(h)\gamma_{jj}(h) \geq |\gamma_{ij}(h)|^2$ (21)



La función de variograma cruzado se relaciona con la covarianza cruzada a través de la siguiente ecuación:

$$\gamma_{ij}(h) = C_{ij}(0) - \frac{1}{2} [C_{ij}(h) + C_{ji}(h)]\tag{22}$$

Para el caso del variograma cruzado, este necesita tener los datos de las diferentes variables medidas en los mismos sitios [16], por lo que no puede ser utilizado para el caso de un muestreo heterotópico total. En tal contexto, se puede utilizar la covarianza cruzada, o en la situación que se necesite usar el variograma cruzado y exista muestreo heterotópico, se cuenta con el seudo variograma cruzado.

2.1.1.5 Modelo lineal de correogionalización

El cokriging, al igual que el kriging, utiliza un sistema de ecuaciones basadas en el variograma (en este caso simple y cruzado), por lo que es necesario contar con una función que modele el

variograma experimental. En el caso univariable, se define un variograma anidado de una sola variable donde se suman varios variogramas elementales. En el caso multivariable, es necesario modelar un conjunto de variogramas simples y cruzados. Las combinaciones de estos se conocen como variograma modelado multivariable o modelo de correogionalización.

El modelo lineal de correogionalización (MLC) plantea, como lo dice su nombre, un modelo lineal, implicando que los variogramas simples y cruzados pueden ser expresados como una suma de modelos básicos (efecto pepita, esférico, exponencial, Gaussiano, seno cardinal, potencia, entre otros) idénticos, generalizando el concepto de anidado para el caso en que se trabaje con múltiples variables. Para N variables, el modelo se define como sigue:

$$\gamma_{ij}(h) = \sum_{u=1}^S b_{ij}^u g_u(h), \quad \forall i, j \in [1, N], \quad (23)$$

y matricialmente,

$$\Gamma(h) = \sum_{u=1}^S \mathbf{B}_u g_u(h), \quad (24)$$

donde $\Gamma(h) = [\gamma_{ij}(h)]_{i,j=1,\dots,N}$ es la matriz de los variogramas simples y cruzados, $g_u(h)$ es un modelo básico de variograma elegido y $\mathbf{B}_u = [b_{ij}^u]_{i,j=1,\dots,N}$ se denomina matriz de correogionalización.

El ajuste de los parámetros del modelo anidado se hace mediante un procedimiento de mínimos cuadrados (para el caso multivariable) [17]. Una de las condiciones para que el modelo sea válido, es que cada $\mathbf{B}_u \{ \forall u = 1, \dots, S \}$ sea simétrica de tipo positivo. Esto se comprueba al verificar que todos los valores propios de dicha matriz sean mayores que cero, ya que esta debe ser invertible.

2.1.2 DISEÑO DE LA VECINDAD DE COKRIGING

La vecindad del cokriging se refiere al dominio en donde se va a realizar la estimación multivariable. Esta puede utilizar todos los datos o sólo algunos de ellos, lo que corresponde a ser una

vecindad única o una vecindad móvil, respectivamente. En este caso, la selección de datos resulta ser más compleja que para una estimación univariable, por lo que se debe considerar el tipo de muestreo que se realiza, así como el modelo de correogionalización. Si las variables tienen correlación donde sus variogramas simples y cruzados son proporcionales entre sí, y el muestreo además es homotópico³, las covariables no estarían aportando información. Siguiendo en la misma situación, pero con un muestreo heterotópico parcial⁴, las covariables sí aportan información en los sitios donde no se cuentan con datos de la variable primaria, por lo que resultaría favorable utilizar una vecindad “dislocada”, descartando los datos secundarios colocalizados con datos primarios [15].

Buscar los puntos con información que estén más cercanos al sitio a estimar o realizar una búsqueda para los datos de la variable a estimar y otra para las variables secundarias, son las estrategias más comunes de búsqueda. Sin embargo, la primera estrategia resulta ser más fácil de aplicar ya que permite estimar todas las variables en forma simultánea. Por el contrario, en la segunda estrategia se deben resolver múltiples veces las ecuaciones de cokriging, por lo que su eficacia depende del número de variables.

2.1.3 HIPÓTESIS DEL COKRIGING ORDINARIO

Sean un conjunto de variables aleatorias modeladas por funciones aleatorias Z_1, \dots, Z_N estacionarias, de medias $\{m_i$ con $i = 1, \dots, N\}$ desconocidas y covarianzas (o varianzas) simples $C(h)$ y cruzadas $\{C_{ij}(h)$ con $i, j = 1, \dots, N\}$ conocidas, donde se supone que se quiere estimar la variable Z_1 en el sitio x_0 del espacio.

2.1.4 PRINCIPIOS DEL COKRIGING ORDINARIO

Para construir el sistema de cokriging ordinario se deben aplicar tres restricciones necesarias de linealidad, insesgo y optimalidad [15, 18-19].

³ Tipo de muestreo donde se conocen las variables en todos los sitios estudiados.

⁴ Tipo de muestreo en el que solo una parte de los sitios de medición son comunes en todas las variables.

1. Restricción de linealidad

La primera restricción consiste en escribir al estimador $Z_1^*(x_0)$ como una combinación lineal ponderada de los datos disponibles de la variable primaria y las diferentes variables secundarias en la vecindad de x_0 . Cada una de las variables es definida en un conjunto de muestras de (posiblemente diferente) tamaño n_i y el estimador es definido como:

$$Z_1^*(x_0) = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i Z_i(x_{\alpha}^i), \quad (25)$$

donde x_{α}^i representa al α -ésimo punto con dato de la i -ésima variable Z_i y $\lambda_{\alpha}^i \{ \forall i = 1, \dots, N, \forall \alpha = 1, \dots, n_i \}$ representa los ponderadores de las variables observadas, las que son incógnitas en el problema.



2. Restricción de insesgo

La segunda restricción exige que el error de estimación tenga una esperanza nula, es decir:

$$E[Z_1^*(x_0) - Z_1(x_0)] = 0 \quad (26)$$

Esta condición se satisface eligiendo pesos cuya suma resulte uno para la variable de interés y cero para las variables auxiliares.

3. Restricción de optimalidad

Finalmente, la última restricción requiere que la varianza del error de estimación sea mínima, la que se resuelve utilizando el método de Lagrange. La restricción es expresada en función de las covarianzas simples y cruzadas:

$$\min\{\text{var}[Z_1^*(x_0) - Z_1(x_0)]\}, \quad (27)$$

con:

$$\begin{aligned}
 \text{var}[Z_1^*(x_0) - Z_1(x_0)] &= \sum_{i=1}^N \sum_{j=1}^N \sum_{\alpha=1}^{n_i} \sum_{\beta=1}^{n_j} \lambda_{\alpha}^i \lambda_{\beta}^j C_{ij} (x_{\alpha}^i - x_{\beta}^j) \\
 &\quad - 2 \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i C_{i1} (x_{\alpha}^i - x_0) + C_{11}(0) \\
 &\quad + 2\mu_1 \left(\sum_{\alpha=1}^{n_1} \lambda_{\alpha}^1 - 1 \right) + 2 \sum_{i=2}^N \mu_i \left(\sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i \right),
 \end{aligned} \tag{28}$$

donde μ_1, \dots, μ_N son incógnitas que representan a los multiplicadores de Lagrange.

2.1.5 DETERMINACIÓN DEL ESTIMADOR

La aplicación de las restricciones de ingreso y optimalidad conducen al siguiente sistema de ecuaciones, cuya resolución entrega los ponderadores del cokriging (incógnitas de la ecuación (25)) y multiplicadores de Lagrange:

$$\left\{ \begin{aligned}
 \sum_{j=1}^N \sum_{\beta=1}^{n_j} \lambda_{\beta}^j C_{ij} (x_{\alpha}^i - x_{\beta}^j) + \mu_i &= C_{i1} (x_{\alpha}^i - x_0), \quad \forall i = 1, \dots, N, \forall \alpha = 1, \dots, n_i \\
 \sum_{\alpha=1}^{n_1} \lambda_{\alpha}^1 &= 1 \\
 \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i &= 0, \quad \forall i = 2, \dots, N
 \end{aligned} \right. \tag{29}$$

2.1.6 VARIANZA DEL COKRIGING ORDINARIO

La varianza mínima del error de estimación del cokriging ordinario en el sitio x_0 , se expresa de la siguiente manera:

$$\sigma_{CKO-Z_1}^2(x_0) = C_{11}(0) - \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{\alpha}^i C_{i1}(x_{\alpha}^i - x_0) - \mu_1 \quad (30)$$

2.2 REDES NEURONALES

Las redes neuronales artificiales (ANN por sus siglas en inglés) son definidas como un conjunto de algoritmos matemáticos que procesan información, que surgen como un intento de reproducir el funcionamiento de las neuronas en el cerebro [20]. Las ventajas de utilizar ANN son el aprendizaje adaptativo, la autoorganización, la tolerancia a fallos, la operación en tiempo real, la fácil inserción dentro de la tecnología existente, entre otras [21].

2.2.1 NOCIÓN DE NEURONA BIOLÓGICA

Para entender cómo operan las redes neuronales, se debe tener conocimiento de cómo funcionan las neuronas en el cerebro. Una neurona biológica es una célula del sistema nervioso, que se compone de un cuerpo celular (llamado soma) y varias extensiones desde este, conocidas como dendritas y axón [21]. En términos generales, las dendritas con el cuerpo celular reciben y procesan la información y, si se activa el potencial de acción, emite una señal. El axón transporta estas señales a sus terminales, los que permiten conectarse con otras neuronas para el traspaso de información. Esta conexión se denomina sinapsis y es unidireccional.

2.2.2 NEURONA ARTIFICIAL

A partir del funcionamiento de la neurona biológica, nace un modelo de neurona artificial (Figura 2) [20]. Al igual que la neurona biológica, la neurona artificial (o nodo) recibe entradas provenientes de otras neuronas o células, donde el vector de entradas es $\mathbf{x} = [x_1, x_2, \dots, x_N]$. Esta información recibida se modifica por un vector w_{ji} que representa a los pesos sinápticos, el cual determina el efecto que la unidad de entrada i tiene en la unidad de salida j , simulando la sinapsis. El parámetro θ_j simboliza el umbral de una neurona. La entrada neta Net_j es la sumatoria entre las diferentes entradas, modificadas por los pesos sinápticos, donde se incluye su umbral θ_j ,

representando el estímulo total de la neurona. Posterior a eso, la activación o no de ésta depende de la llamada función de activación, la cual es utilizada para poder aplicar las ANN a diversos problemas reales, donde la entrada neta se evalúa en esta función obteniéndose la salida y_j de la red. Finalmente, la salida y_j se puede propagar hacia otras neuronas o puede ser el resultado final de la red.

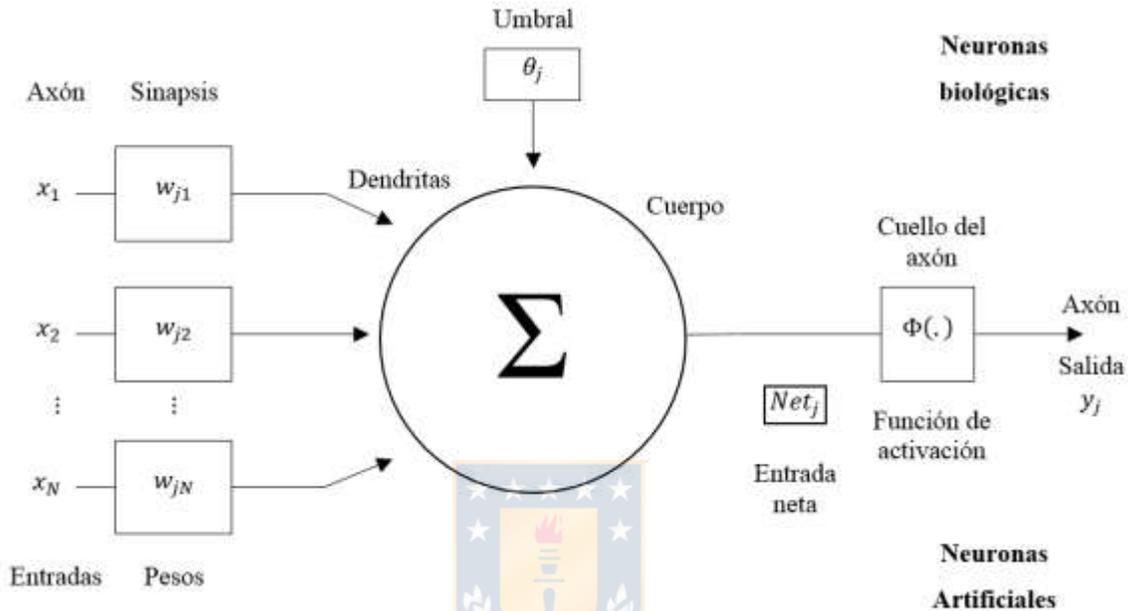


Figura 2. Modelo de neurona artificial.

El cálculo de la entrada neta se representa en la siguiente ecuación:

$$Net_j = \sum_{i=1}^N x_i w_{ji} - \theta_j \quad (31)$$

Por otro lado, el cálculo de la salida y_j , determinada por la función de activación $Fact_j$, se expresa como:

$$y_j = Fact_j(Net_j), \quad (32)$$

donde la función de activación es típicamente del tipo lineal, escalón, sigmoideal o gaussiana [20]. En el Anexo A. 1 se encuentran las descripciones de las funciones de activación más utilizadas.

2.2.3 RED NEURONAL ARTIFICIAL

Por si sola, la neurona artificial carece de capacidad de procesamiento y aplicabilidad. Es por ello que nacen las redes neuronales artificiales, pues la combinación de ellas puede resolver problemas muy complejos [21]. Las redes neuronales son un sistema de procesamiento de información donde su funcionamiento se inspira en imitar a las redes neuronales biológicas. En la red, la distribución de las neuronas se realiza formando capas o niveles, donde cada capa tiene un número de neuronas determinado. Estas capas se caracterizan en que cada neurona perteneciente a ella recibe el mismo tipo de información. En una ANN se diferencian tres tipos de capas:

- **Capas de entrada:** son aquellas que reciben las variables de entradas, es decir, reciben la información de las fuentes externas.
- **Capas ocultas:** son capas internas de la red que no son observables directamente desde el exterior. Las señales de entrada y salida de las neuronas pertenecientes a estas capas permanecen dentro de la red, donde el número de capas pueden ser mayor o igual a cero. Las neuronas de estas capas se pueden interconectar de distintas maneras, determinando, junto al número de capas, la estructura de la red neuronal.
- **Capas de salida:** son aquellas donde el conjunto de neuronas transfiere la información procesada hacia el exterior.

En la siguiente figura se puede apreciar la arquitectura de una ANN con varias capas.

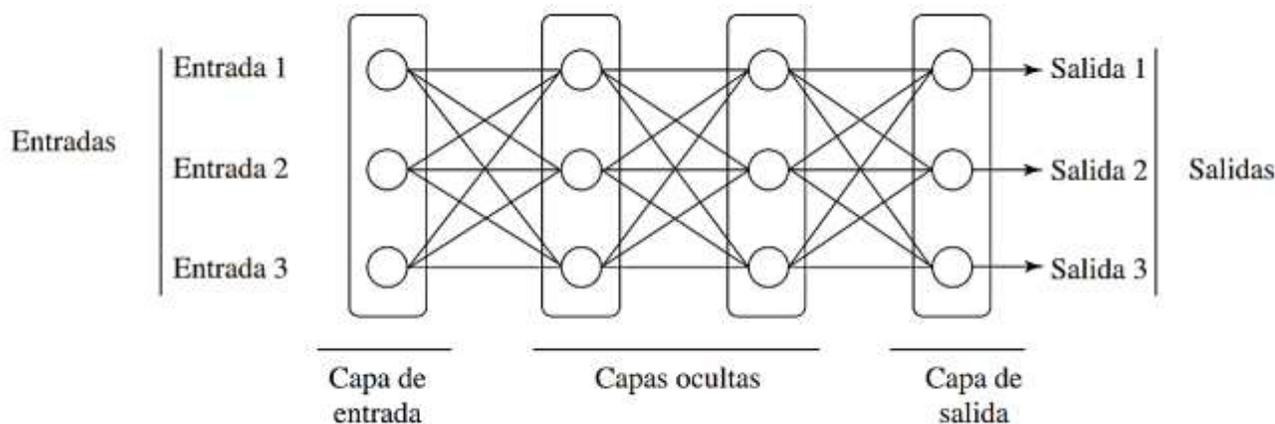


Figura 3. Modelo de red neuronal artificial con las conexiones hacia adelante.

2.2.4 ESTRUCTURA DE UNA ANN

La estructura de una ANN está determinada en cómo estén organizadas las neuronas, es decir, en la cantidad de capas y en cómo se conecten las neuronas, y eso se relaciona directamente con el algoritmo de aprendizaje utilizado para entrenar la red [20]. Si se considera el número de capas, se habla de redes monocapa o redes multicapas, mientras que si se habla del flujo de datos puede tratarse de redes unidireccionales o de propagación hacia delante (*feedforward*) o redes de propagación hacia atrás (*feedback*).

2.2.4.1 Según número de capas

- **Redes monocapa:** cuando una ANN cuenta con un vector de entradas y solo una capa de procesamiento, siendo el caso más simple, quiere decir que no cuenta con capas ocultas, por lo que la red se denomina monocapa. Este tipo de redes típicamente se utiliza para problemas de autoasociación y de *clustering* [22].
- **Redes multicapa:** son aquellas en donde las neuronas están organizadas en más de una capa, contando con una o más capas ocultas.

2.2.4.2 Según formas de conexión entre neuronas

- **Redes *feedforward*:** es cuando las conexiones entre neuronas fluyen en un único sentido hacia delante (no existen conexiones hacia atrás), desde las neuronas de entrada a la capa o capas de procesamiento, hasta llegar a la capa de salida (Figura 4).

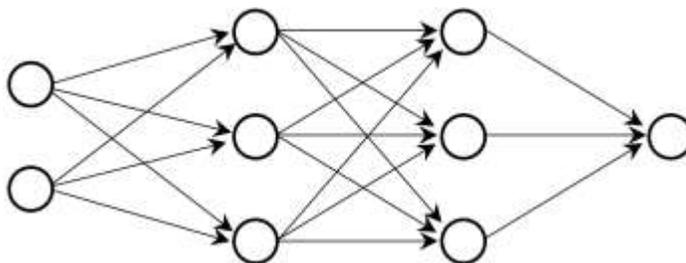


Figura 4. Ejemplo de red *feedforward*.

- **Redes *feedback*:** es cuando existen conexiones hacia atrás (ver Figura 5), donde la conexión puede ser entre una misma neurona (a), entre neuronas de una misma capa (b) y entre neuronas de una capa a una capa anterior (c).

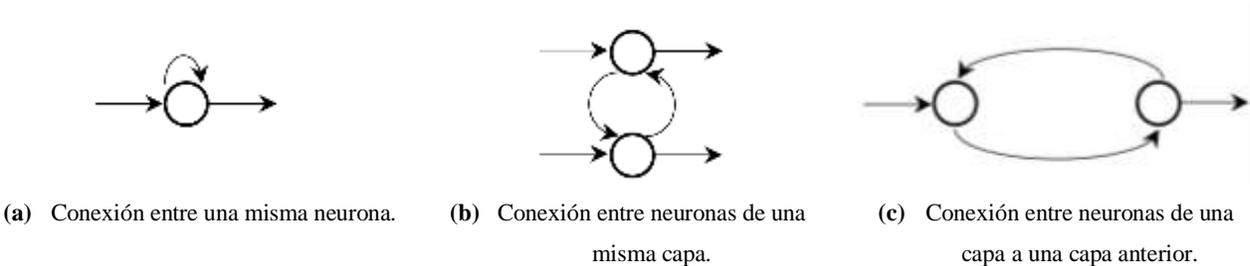


Figura 5. Ejemplos de red *feedback*.

2.2.5 APRENDIZAJE EN LAS ANN

Las ANN son entrenadas para aprender a calcular la salida correcta para cada vector de entrada, donde, para lograr este fin, se utiliza un conjunto de datos de entrenamiento. Como las entradas y salidas no pueden cambiar, lo que se modifica son los pesos sinápticos. Este proceso termina cuando los valores de los pesos permanecen estables. Los cambios en los pesos sinápticos se reducen a creación y/o destrucción de las neuronas en la red. En general, se consideran dos tipos de reglas, las que se conocen como aprendizaje supervisado y aprendizaje no supervisado.

La siguiente ecuación es una generalización de los cambios en los pesos sinápticos:

$$w_{ji}(t + 1) = w_{ji}(t) + \Delta w_{ji}(t), \quad (33)$$

donde t se refiere a la etapa de aprendizaje, $w_{ji}(t + 1)$ es el peso nuevo y $w_{ji}(t)$ es el peso viejo.

Otro criterio que se puede utilizar para diferenciar las reglas de aprendizaje es el aprendizaje *on line* y aprendizaje *off line*. En el primer caso, la red aprende durante su funcionamiento habitual, mientras que para el segundo caso se supone la desconexión de la red hasta que el proceso de aprendizaje termine [21].

2.2.5.1 Redes con aprendizaje supervisado

El aprendizaje supervisado es aquel proceso de entrenamiento de la red neuronal controlado por un agente externo (supervisor o maestro). Este supervisor conoce las salidas deseadas de la red correspondiente a cada entrada, la cual es comparada con la salida computada. Si las salidas no coinciden, se procede a modificar los pesos sinápticos, con el fin de obtener una salida lo más próxima a la deseada.

Existen tres formas de llevar a cabo el aprendizaje supervisado [22]:

- **Aprendizaje por corrección de error:** el aprendizaje por corrección de error consiste en ajustar los pesos sinápticos en función de la diferencia entre las salidas obtenidas y las salidas deseadas, siendo este el error cometido. La siguiente regla es una de las más sencillas para este tipo de aprendizaje:

$$\Delta w_{ji} = \alpha x_i (d_j - y_j), \quad (34)$$

donde $\Delta w_{ji} = w_{ji \text{ actual}} - w_{ji \text{ anterior}}$ es la variación en el peso sináptico entre las neuronas i y j , x_i es el valor de la i -ésima entrada a la j -ésima neurona, d_j es el valor de salida deseado por la neurona j , y_j es el valor de salida obtenido por la neurona j y $\alpha \{0 < \alpha \leq 1\}$ es un factor que regula la velocidad de aprendizaje. La regla anteriormente descrita presenta algunas limitaciones, como el no considerar la magnitud del error global cometido durante el proceso completo de aprendizaje. Sin embargo, más adelante fue desarrollado un algoritmo de aprendizaje más eficiente conocido como algoritmo de retropropagación del error o *backpropagation*.

- **Aprendizaje por refuerzo:** este aprendizaje es más lento que el anteriormente descrito, ya que se basa en que, durante el entrenamiento, el supervisor no proporcione exactamente la salida deseada, sino que indique, mediante una señal de refuerzo, si la salida obtenida se ajusta a la salida deseada (éxito = 1 o fracaso = -1). De esa manera, los pesos se ajustan basándose en un mecanismo de probabilidades.
- **Aprendizaje estocástico:** este tipo de aprendizaje se basa principalmente en modificar los pesos sinápticos aleatoriamente, evaluando su efecto en términos de la salida deseada y de una distribución de probabilidad. Se suele hacer una analogía en términos termodinámicos, donde la

red se asocia con un sólido físico que tiene cierto estado energético. Este representa la estabilidad de la red, de tal forma que el estado de mínima energía ocurre cuando los pesos consiguen ser los adecuados para lograr la salida deseada. De esa manera, el aprendizaje consiste en modificar aleatoriamente los pesos y determinar la energía de la red. Si esta es menor luego de la modificación, se acepta el cambio, pues el comportamiento de la red se acerca al deseado. En caso contrario, se acepta el cambio en función de una distribución de probabilidades.

2.2.5.2 Redes con aprendizaje no supervisado

El aprendizaje no supervisado es cuando el vector de datos que describe el problema se presenta directamente a la red, sin un supervisor que guíe el aprendizaje. El algoritmo es el que modifica los pesos sinápticos con el fin de que las salidas sean las deseadas, es decir que, a entradas similares, la red compute la misma salida. Así, las salidas están asociadas a las entradas según el proceso de entrenamiento.

En general, se suelen considerar dos tipos de algoritmos de aprendizaje no supervisado [22]:

- **Aprendizaje Hebbiano:** el aprendizaje hebbiano tiene por objetivo medir la familiaridad o extraer características de los datos de entrada. Se basa en suponer que si dos neuronas N_i y N_j toman el mismo estado (ambas activas o inactivas), el peso sináptico entre ambas incrementa. Se utiliza en neuronas binarias, donde las entradas y salidas son $\{-1,1\}$ o $\{0,1\}$.
- **Aprendizaje competitivo y cooperativo:** en esta clase de aprendizaje, se dice que las neuronas compiten (y cooperan) unas con otras para llevar un objetivo a cabo. Lo que se pretende en este aprendizaje es que cuando ingrese una cierta información a la red, sólo una neurona de salida, o una por grupo, se active. Entonces, el aprendizaje se fundamenta en la clasificación o agrupamiento de los datos de entrada. La información similar se clasifica para formar parte de una misma categoría, por lo que deben activar a la misma neurona de salida.

En el Anexo A. 2 se encuentra la descripción de los tipos de ANN más utilizados.

2.2.6 ALGORITMO BACKPROPAGATION

Uno de los modelos de redes neuronales más populares y utilizados es el *backpropagation* por su capacidad de solucionar problemas de alta complejidad. Este modelo está diseñado para entrenar una red multicapa con conectividad total (es decir, que este completamente interconectada entre capas, como en la Figura 3), con propagación hacia delante y aprendizaje supervisado [23].

El algoritmo *backpropagation* se basa en propagar el error de la capa de salida hacia atrás, ya que este es el único que se puede calcular de forma exacta, y así estimar el error en las salidas de las neuronas en las capas ocultas. Este proceso se repite hasta que todas las neuronas de las capas ocultas tengan el error que aportaron relativo al error total, con el objetivo de modificar sus pesos sinápticos para que en una próxima instancia la salida sea más cercana a la deseada [20].

El algoritmo se detiene una vez que su condición de parada sea verificada, como que el error de la salida sea reducido a un valor permitido, o el número de iteraciones propuesto haya sido superado.

2.2.6.1 Pasos del algoritmo *backpropagation*

A continuación, se presentan los pasos que sigue el algoritmo *backpropagation* [23].

1. Aplicar el vector de entrada $\mathbf{x}_p = (x_{p1}, x_{p2}, \dots, x_{pN})^t$ a las unidades de entrada.
2. Calcular los valores de las entradas netas para las unidades de la capa oculta:

$$\text{Net}_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h, \quad (35)$$

donde N es la dimensión del vector de entrada, w_{ji}^h es el peso de interconexión entre la neurona i -ésima de la entrada y la j -ésima de la capa oculta y θ_j^h es el umbral de la j -ésima neurona de la capa oculta. El índice «h» se refiere a magnitudes de la capa oculta.

3. Calcular las salidas de la capa oculta:

$$i_{pj} = f_j^h(\text{Net}_{pj}^h), \quad (36)$$

donde f_j^h es la función de activación de la j -ésima neurona de la capa oculta.

4. Calcular los valores netos de las entradas para las unidades de la capa de salida:

$$\text{Net}_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + \theta_k^o, \quad (37)$$

donde L es el número de neuronas de la capa oculta h, w_{kj}^o es el peso de la interconexión entre la j -ésima neurona de la capa oculta y la k -ésima neurona de la capa de salida y θ_k^o es el umbral de la k -ésima neurona de la capa de salida. El índice «o» se refiere a magnitudes de la capa de salida.

5. Calcular las salidas de la red:

$$o_{pk} = f_k^o(\text{Net}_{pk}^o), \quad (38)$$

donde f_k^o es la función de activación de la k -ésima neurona de la capa de salida.

6. Calcular los términos de error para las unidades de salida:

$$\delta_{pk}^o = (y_{pk} - o_{pk}) f_k^{o'}(\text{Net}_{pk}^o), \quad (39)$$

donde y_{pk} es el valor de salida deseado para la k -ésima neurona de la capa de salida.

7. Calcular los términos de error para las unidades ocultas:

$$\delta_{pj}^h = f_j^{h'}(\text{Net}_{pj}^h) \sum_{k=1}^M \delta_{pk}^o w_{kj}^o, \quad (40)$$

donde M es el número de neuronas de la capa de salida, es decir, la dimensión del vector de salida.

8. Actualizar los pesos de la capa de salida:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta \delta_{pk}^o i_{pj}, \quad (41)$$

donde η es la tasa o velocidad de aprendizaje.

9. Actualizar los pesos de la capa oculta:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta \delta_{pj}^h x_{pi} \quad (42)$$

10. Verificar si el error global cumple con la condición de finalizar [20]:

$$E_p = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^M (y_{pk} - o_{pk})^2, \quad (43)$$

donde P es el número de patrones de entrenamiento (información que debe aprender la red).

2.3 ÁRBOLES DE REGRESIÓN

Un árbol de decisión es una técnica predictiva que consiste en una división jerárquica y secuencial, donde cada una describe posibles decisiones que llevan a ciertos resultados. Estos comienzan generalmente con un único nodo y luego este se ramifica en distintos resultados, donde

cada resultado crea un nodo que puede o no seguir ramificándose, dando una forma semejante a la de un árbol.

Un árbol de decisión puede ser utilizado para crear modelos predictivos, con el fin de ser empleados en el aprendizaje automático, la minería de datos y las estadísticas. Este método considera las observaciones sobre cierta variable para predecir su valor.

2.3.1 MÉTODO CART

Los árboles de clasificación y regresión, más conocido como método CART (*Classification and Regression Trees*), fueron desarrollados por primera vez por Leo Breiman y otros en 1984 [24]. Se refiere a una metodología en donde se utilizan datos históricos para construir árboles de clasificación o de regresión, ya sea para clasificar o predecir nuevos datos [25], respectivamente. En estos árboles es posible manipular variables categóricas y/o continuas, según sea el caso. Una de sus grandes ventajas es su robustez a *outliers*, la invarianza en la estructura de sus árboles a transformaciones monótonas de las variables independientes y, en particular, su interpretabilidad.

El árbol de decisión CART es un procedimiento de particionamiento recursivo binario, donde los datos son manejados en su forma “bruta”, es decir, no se requiere ni se recomienda hacer discretización [26]. Se trata de árboles de clasificación cuando la variable dependiente es cualitativa y árboles de regresión cuando la variable dependiente es cuantitativa. La finalidad es obtener una predicción para la clase de pertenencia, en el caso de los árboles de clasificación, o para el valor que toman los individuos que cumplen con las condiciones exigidas en los distintos nodos, para el caso de árboles de regresión.

Se comienzan en un nodo raíz, donde los datos se dividen en dos hijos, y cada uno de ellos se divide en nietos. Los árboles son llevados hasta un tamaño máximo (árbol saturado), sin el uso de una regla de detención, es decir, el proceso de crecimiento continua hasta que no es posible realizar más divisiones debido a la falta de datos y luego este se recorta mediante la poda de costo-complejidad. El árbol de “tamaño correcto” se identifica al evaluar el rendimiento predictivo de cada árbol en la secuencia de poda, el cual se mide con datos de prueba independientes (o por validación cruzada).

2.3.2 ELEMENTOS DE UN ÁRBOL

Existen tres tipos de nodos [27] en los árboles de decisión, ilustrados en la Figura 6:

- **Nodo raíz:** es el primer nodo en la cima, el cual no tiene ramas entrantes y tiene cero o más ramas salientes.
- **Nodos internos:** los nodos internos tienen una rama entrante y dos o más ramas salientes.
- **Nodos terminales:** también llamados hojas, son aquellos que tienen una rama entrante y cero ramas salientes.

En el método CART, tanto en el nodo raíz como en los nodos internos se realizan particiones que resultan en dos nodos adicionales cada uno, los que son llamados nodos hijos (o ramas) izquierdo y derecho.

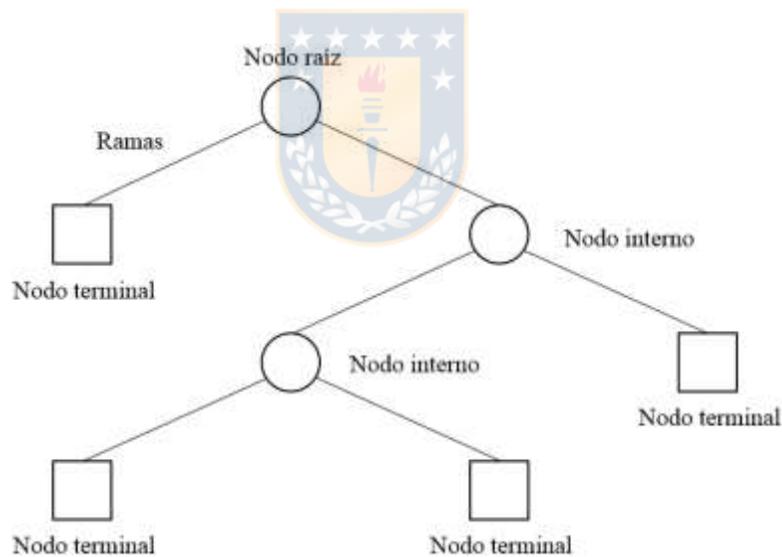


Figura 6. Modelo de árbol de regresión.

2.3.1 PARTICIONAMIENTO RECURSIVO

El particionamiento recursivo se trata de un algoritmo con el cual se construye un árbol de decisión, lo que es imprescindible para el método CART [25].

Sea Y la variable respuesta (también llamada variable dependiente) y sean p las variables predictoras (o variables independientes) x_1, x_2, \dots, x_p , donde las variables x son fijadas e Y es una variable aleatoria. Se requiere establecer una relación entre estas variables con el objetivo de predecir Y a partir de las variables x . En el caso de una variable con respuesta continua, se quiere estimar un funcional de la probabilidad condicional de Y , como lo es la esperanza condicional $E[Y|x_1, x_2, \dots, x_p]$.

El objetivo del particionamiento recursivo es que los nodos hijos sean homogéneos (o puros). Sin embargo, la realidad es que estos nodos son usual y parcialmente puros.

2.3.2 IMPUREZA DEL NODO

La homogeneidad de un nodo se puede medir cuantitativamente mediante la impureza, siendo su indicador el siguiente:

$$\text{Impureza del nodo} = \frac{\text{Número de sujetos que cumplen la característica en el nodo}}{\text{Número total de sujetos en el nodo}} \quad (44)$$

La impureza toma el valor 0 cuando la característica en el nodo no se cumple, y el valor 1 cuando la característica se cumple 100%, es decir, el límite del cociente de la ecuación anterior toma valores 0 o 1 mientras más homogéneo sea el nodo. En el caso de que la respuesta sea continua, una medida de la impureza para un nodo τ es la varianza de la respuesta dentro del nodo:

$$i(\tau) = \sum_{\text{sujeto } i \in \tau} (Y_i - \bar{Y}(\tau))^2, \quad (45)$$

donde $\bar{Y}(\tau)$ es el promedio de las variables Y_i dentro del nodo τ .

2.3.3 REGLAS DE DIVISIÓN DE UN NODO

Las reglas de división del método CART son redactadas de la forma [26]:

“Una instancia va a la izquierda si se cumple CONDICIÓN, de lo contrario va a la derecha”,

donde la condición es expresada como $X_i \leq C$ para atributos continuos. Para hacer una división homogénea, los rangos seleccionados de todas las variables predictoras deben ser tal que el valor del cociente de la ecuación (44) sea lo más próximo a 0 o 1 [25].

Antes de seleccionar la mejor división de un nodo τ en dos nodos hijos, τ_L y τ_R , es necesario definir y utilizar la bondad de una división s :

$$\Delta I(\tau) = i(\tau) - i(\tau_L) - i(\tau_R), \quad (46)$$

donde τ es el nodo padre del nodo izquierdo τ_L y del nodo derecho τ_R . Lo que hace la ecuación anterior es medir el grado de reducción de la impureza cuando se pasa del nodo padre a los nodos hijos, donde la división s es tal que $\Delta I(\tau)$ sea máxima.

2.3.4 PODADO DEL ÁRBOL (*PRUNING*)

El procedimiento que se realiza en el particionamiento recursivo prosigue hasta lograr que el árbol sea saturado. Esto ocurre cuando los nodos descendientes no pueden seguir dividiéndose, como lo es el caso en que en un nodo solo quede un sujeto. Si un nodo no es dividido, pasa a ser un nodo terminal.

En general, un árbol saturado suele ser bastante grande como para ser utilizado, pues con los nodos terminales muy pequeños no es posible hacer una inferencia estadística razonable, esto debido a que existe un sobreajuste. Es en este contexto en que se utiliza la poda del árbol saturado, que básicamente consiste en encontrar un subárbol de este que sea el más predictivo posible y lo menos vulnerable al ruido de los datos, la cual se realiza desde el último nivel del árbol hacia arriba.

En otras palabras, lo que se quiere definir es una medida de calidad del árbol, la que depende de la calidad de sus nodos terminales. Para un árbol de regresión T se hace uso de $i(\tau)$ para definir el costo del árbol:

$$R(T) = \sum_{\tau \in \tilde{T}} i(\tau), \quad (47)$$

donde \tilde{T} es el conjunto de nodos terminales de T . De un árbol saturado inicialmente T_0 , se quiere seleccionar al mejor subárbol T^* , tal que $R(T)$ sea mínimo.

2.3.5 COSTO-COMPLEJIDAD

Como ya ha sido mencionado, un árbol de grandes dimensiones puede tener problemas de sobreajuste, por lo que el tamaño del árbol es importante cuando se dan conclusiones sobre la muestra o población estudiada. Entonces, para definir la calidad del árbol se debe tener en consideración tanto la calidad de los nodos terminales como el tamaño del árbol, traducidos en el número de nodos terminales. Es por ello que se define el costo-complejidad del árbol T como:

$$R_\alpha(T) = R(T) + \alpha |\tilde{T}|, \quad (48)$$

donde $\alpha (\geq 0)$ es el parámetro de complejidad y $|\tilde{T}|$ es el número de nodos terminales en T llamado complejidad del árbol T . La principal diferencia entre la ecuación (47) y (48), es que $R_\alpha(T)$ penaliza a un árbol de gran tamaño.

2.4 SUPPORT VECTOR MACHINE

El concepto de Máquinas de Soporte Vectorial (SVM, del inglés *Support Vector Machine*) fue introducido por Vladimir Vapnik y sus colaboradores en 1995 [28]. Se trata de un método basado en la teoría del aprendizaje estadístico, que puede ser utilizado tanto en clasificación como en regresión, donde en este último caso, las SVM son comúnmente llamadas *Support Vector Regression* (SVR). Su formulación se basa en el principio *Structural Risk Minimization* (SRM) [28] o Minimización del Riesgo Estructural, donde, a diferencia del principio de Minimización del Riesgo Empírico (ERM) que sugiere que se debe minimizar el riesgo empírico a cualquier costo [29], minimiza el riesgo funcional con respecto al riesgo empírico y a la dimensión VC (Vapnik-Chervonenkis) de un set de

funciones [30]. En otras palabras, el principio SRM encuentra la función que, para una cantidad fija de datos, alcanza el mínimo del riesgo garantizado [29].

El trabajo de las SVR es obtener un hiperplano que mejor represente al conjunto de entrenamiento, donde este puede ser tanto lineal como no lineal. Una vez obtenido el hiperplano, se construyen unas “bandas” paralelas a este cubriendo la mayor cantidad de datos, las que son llamadas vectores de apoyo o de soporte. Los datos que no son cubiertos por los vectores de soporte son los errores y estos deben ser considerados en la fórmula final del algoritmo.

En las SVR, el objetivo es encontrar una función $f(x)$ que tenga una desviación máxima ε de los datos realmente obtenidos y_i y, al mismo tiempo, que sea lo más “plana” posible. En otras palabras, no importan los errores mientras sean menores que ε , pero no se acepta ninguna desviación mayor que esta [31].

Algunas de las ventajas de las SVM es que el modelo suele ser sencillo, robusto y tiene una buena generalización para nuevos datos cuando el modelo está bien parametrizado.

2.4.1 SVR CASO LINEAL

Sea un conjunto de ejemplos de entrenamiento $S = \{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathcal{X} \times \mathbb{R}$, donde \mathcal{X} denota el espacio de los datos de entrada (por ejemplo, $\mathcal{X} = \mathbb{R}^d$), siendo asumido que los valores de y_i de todos los ejemplos de S se pueden ajustar (o cuasi-ajustar) por medio de una función lineal. El objetivo es encontrar los parámetros $w = (w_1, \dots, w_d)$ y b tal que la función lineal pueda ser definida como [31]:

$$f(x) = \langle w, x \rangle + b, \quad (49)$$

donde $w \in \mathcal{X}$, $b \in \mathbb{R}$ y $\langle \cdot, \cdot \rangle$ denota el producto punto en \mathcal{X} . En la práctica, lograr que los ejemplos de entrenamiento se ajusten con un error cero al modelo lineal es difícil, por lo que se acude al concepto llamado margen blando, aceptando cierto error entre el valor predicho y el valor real. Así, se emplea la función de pérdida ε -insensible, L_ε (ver Figura 7), la cual es una función lineal con una zona insensible, de anchura 2ε , en la que el error es nulo, descrita por [32]:

$$L_\varepsilon(y, f(x)) = |\xi|_\varepsilon = \begin{cases} 0 & \text{si } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{en otro caso} \end{cases} \quad (50)$$

La función anterior se utiliza para permitir cierta dispersión en el regresor, de modo que los ejemplos de entrenamiento que queden confinados en la región tubular, definida por $\pm\varepsilon$, no son considerados vectores de soporte, reduciendo el número de estos.

“Planeidad”, en el caso de la ecuación (49), significa buscar un w pequeño. Una forma de asegurarse de que esto ocurra es minimizando la norma vectorial de este, es decir, $\|w\|^2 = \langle w, w \rangle$. Lo anterior se puede escribir como un problema de optimización convexa [31]:

$$\begin{aligned} &\text{Minimizar} && \frac{1}{2} \|w\|^2 \\ &\text{Sujeto a} && \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \end{aligned} \quad (51)$$

La suposición anterior es que la función f existe tal que aproxima todos los pares (x_i, y_i) con una precisión ε , es decir, que el problema de optimización convexa es factible. En algunas ocasiones este puede no ser el caso, o también se pueden permitir algunos errores. Se pueden introducir variables de holgura ξ_i y ξ_i^* que permiten cuantificar la magnitud de dicho error (ver Figura 7), donde la formulación del problema primal queda como:

$$\begin{aligned} &\text{Minimizar} && \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ &\text{Sujeto a} && \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad \forall i = 1, \dots, n, \end{cases} \end{aligned} \quad (52)$$

donde la constante $C > 0$ es un parámetro de regularización, que determina la planeidad de f y la cantidad hasta la cual se toleran desviaciones mayores que ε .

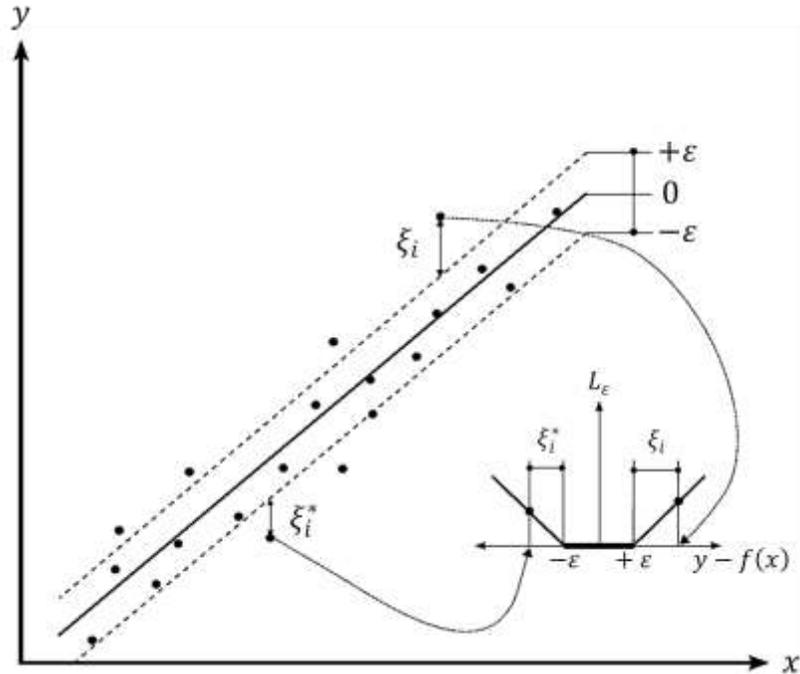


Figura 7. SVR con margen blando.

2.4.1.1 Problema dual

Las SVM abordan problemas de optimización con restricciones lineales, los que se resuelven utilizando la Teoría de la Optimización [32] (ver Anexo B. 1).

Para realizar la transformación del problema primal al problema dual se realizan una serie de pasos:

- **Paso 1:** Obtención de la función lagrangiana.
- **Paso 2:** Aplicación de las condiciones de KKT.
- **Paso 3:** Establecer las relaciones entre las variables del problema primal con las del problema dual.
- **Paso 4:** Establecer restricciones adicionales de las variables duales.
- **Paso 5:** Del resultado obtenido del paso 3, eliminar las variables primales de la función lagrangiana.

Los pasos anteriores son desarrollados en el Anexo B. 2. Finalmente, la optimización del problema dual queda de la siguiente manera [31]:

$$\begin{aligned} \text{Maximizar} \quad & -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle \mathbf{x}_i, \mathbf{x}_j \rangle - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \\ \text{Sujeto a} \quad & \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \quad \forall i = 1, \dots, n, \end{cases} \end{aligned} \quad (53)$$

donde α_i y α_i^* son multiplicadores de Lagrange. Luego, el regresor que se asocia a la función lineal se puede reescribir como:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle \mathbf{x}_i, \mathbf{x} \rangle + b \quad (54)$$

2.4.1.2 Obtención del valor de b

La obtención del valor de b se realiza utilizando las condiciones de KKT, las que afirman que, en el punto de la solución, el producto entre las variables duales y las restricciones tiene que desaparecer. El paso a paso para calcular esta variable esta descrito en el Anexo B. 3.

De las condiciones de KKT, algunas conclusiones útiles son las siguientes [33]:

- En primer lugar, solo los ejemplos (x_i, y_i) con $\alpha_i^{(*)} = C$ quedan fuera de la zona tubular ε -insensible, es decir, $\xi_i^{(*)} > 0$.
- En segundo lugar, se tiene que $\alpha_i \alpha_i^* = 0$, significando que nunca puede haber un conjunto de variables duales α_i, α_i^* que sean simultáneamente distintas de cero.
- Por último, para $\alpha_i^{(*)} \in (0, C)$ se tiene que $\xi_i^{(*)} = 0$.

Por lo tanto, en resumen, si $\alpha_i > 0$ y $\alpha_i < C$, b se puede calcular como:

$$b = y_i - \varepsilon - \langle \mathbf{w}, \mathbf{x}_i \rangle, \quad \text{si } 0 < \alpha_i < C \quad (55)$$

Por otro lado, si $\alpha_i^* > 0$ y $\alpha_i^* < C$, el cálculo de b se realiza como:

$$b = y_i + \varepsilon - \langle \mathbf{w}, \mathbf{x}_i \rangle, \quad \text{si } 0 < \alpha_i^* < C \quad (56)$$

Las condiciones asociadas a las ecuaciones (55) y (56) no pueden ser ciertas simultáneamente, por lo que el valor del parámetro b es siempre único. Cabe destacar que existen otras formas de calcular b , como lo es en el contexto de optimización de puntos interiores, resultando ser un subproducto del proceso de optimización (Ver [33]).

2.4.2 SVR CASO NO LINEAL

Existen ejemplos de entrenamiento que no pueden ajustarse por una función lineal. En ese caso, el siguiente paso es hacer el que algoritmo SV sea no lineal, lo que puede lograrse procesando previamente los ejemplos de entrenamiento x_i mediante un mapa $\Phi : \mathcal{X} \rightarrow \mathcal{F}$ en algún espacio de características \mathcal{F} y luego aplicar el algoritmo SVR estándar [31], donde $\Phi(x) = (\Phi_1(x), \dots, \Phi_m(x))$ es una función no lineal. En otras palabras, el espacio original al que pertenecen los ejemplos se transforma en un nuevo espacio en el que el ajuste de los ejemplos de entrenamiento transformados a un regresor lineal sí es posible [32]. Para que lo anterior ocurra, se necesita de una herramienta denominada función kernel.

2.4.2.1 Kernelización de las SVR

Una función kernel es una función $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ que asigna a cada par de elementos del espacio de entrada original \mathcal{X} un valor real correspondiente al producto punto de las imágenes de dichos elementos en un nuevo espacio de características \mathcal{F} .

El algoritmo SV descrito solo depende de productos puntos entre variables x_i . Por ello, es suficiente saber que la función kernel se define como:

$$k(x, x') = \langle \Phi(x), \Phi(x') \rangle, \quad (57)$$

lo que permite reformular el problema de optimización de la siguiente manera:

$$\begin{aligned} \text{Maximizar} \quad & -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, x_j) - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i(\alpha_i - \alpha_i^*) \\ \text{Sujeto a} \quad & \begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \quad \forall i = 1, \dots, n \end{cases} \end{aligned} \quad (58)$$

En este caso, w puede reescribirse como:

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \Phi(x_i) \quad (59)$$

Así, la expansión de la ecuación (54) queda dada por:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (60)$$

Se debe tener en cuenta que ahora el problema de optimización se basa en encontrar la función más plana en el espacio de características \mathcal{F} , y no en el espacio de entrada \mathcal{X} .

2.4.2.2 Condición para kernels

El teorema de Mercer establece que para garantizar que una función continua simétrica $k(x, x')$ en $L_2(\mathcal{X})$ tiene una expansión:

$$k(x, x') = \sum_{k=1}^{\infty} a_k z_k(x) z_k(x'), \quad (61)$$

con coeficientes positivos $a_k > 0$, es decir, que $k(x, x')$ describa un producto punto en algún espacio de características, es necesario y suficiente que la condición:

$$\int_{\mathcal{X}} \int_{\mathcal{X}} k(x, x') f(x) f(x') \partial x \partial x' \geq 0, \quad (62)$$

sea válida para todo $f \in L_2(\mathcal{X})$, \mathcal{X} siendo un subconjunto compacto de \mathbb{R}^d [28].

Cumpléndose el teorema descrito, un kernel puede darse por una combinación lineal de kernels, una integral de kernels, productos entre kernels, entre otros [31], de la siguiente manera:

1. $k(x, x') = c_1 k_1(x, x') + c_2 k_2(x, x')$ (63)

2. $k(x, x') = \int_{\mathcal{X}} s(x, z) s(x', z) \partial z$ (64)

3. $k(x, x') = k_1(x, x') k_2(x, x')$, (65)

donde k_1 y k_2 denotan kernels admisibles, $c_1, c_2 \geq 0$ y $s(x, x')$ es una función en $\mathcal{X} \times \mathcal{X}$.

2.4.2.3 Ejemplos de kernels

En la Tabla 1 se dan algunos ejemplos de kernels con sus respectivas funciones [33, 34] que han propuesto diferentes autores (Schölkopf, Smola, Vapnik, entre otros), donde se asume que $\mathcal{X} \subset \mathbb{R}^d$.

Tabla 1. Ejemplos de funciones kernels.

	Kernel
Lineal	$k(x, x') = \langle x, x' \rangle$
Polinómico homogéneo	$k(x, x') = \langle x, x' \rangle^p, p \in \mathbb{N}$
Gaussiano	$k(x, x') = \exp\left(-\frac{\ x - x'\ ^2}{2\sigma^2}\right), \sigma > 0$
Sigmoidal	$k(x, x') = \tanh(\kappa\langle x, x' \rangle + \vartheta), \kappa > 0, \vartheta < 0$
Polinomio no homogéneo	$k(x, x') = (\langle x, x' \rangle + c)^p, c \geq 0$
B_n-Spline de orden $2n+1$	$k(x, x') = B_{2p+1}(\ x - x'\),$ con $B_n = \otimes_{i=1}^n I_{[-\frac{1}{2}, \frac{1}{2}]}$, donde I_x denota la función del indicador en el conjunto X y \otimes la operación de convolución $\int f(x')g(x' - x)\partial x'$.
Radial Basis Function (RBF)	Los kernel RBF son kernels que pueden ser escritos como: $k(x, x') = f(d(x, x'))$, donde d es una métrica en \mathcal{X} y f es una función en \mathbb{R}_0^+ . Algunos ejemplos son los kernels Gaussianos y B_n -Spline.
Series de Fourier	Considerando un nuevo espacio de características de dimensiones $2N+1$ y definido en el intervalo $[-\frac{\pi}{2}, \frac{\pi}{2}]$: $k(x, x') = \frac{\sin\left(N + \frac{1}{2}\right)(x - x')}{\sin\left(\frac{1}{2}(x - x')\right)}$

2.4.3 FUNCIONES DE COSTO

2.4.3.1 Riesgo funcional

Como en la sección 2.4.1, supóngase que se tiene un conjunto de ejemplos de entrenamiento $S = \{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathcal{X} \times \mathbb{R}$, donde dicho conjunto ha sido elaborado de manera independiente e idénticamente distribuida con alguna distribución de probabilidad $P(x, y)$ desconocida. El objetivo es encontrar una función f que minimice el riesgo esperado $R[f]$, donde la estimación de f se basa solo en los datos empíricos S . El riesgo esperado se define de la siguiente forma:

$$R[f] = \int c(x, y, f(x)) \partial P(x, y), \tag{66}$$

donde $c(x, y, f(x))$ denota la función de costo que determina como se penalizan los errores de estimación. Una posible aproximación consiste en reemplazar la integral por la estimación empírica, obteniendo el llamado riesgo empírico funcional:

$$R_{emp}[f] = \frac{1}{n} \sum_{i=1}^n c(x_i, y_i, f(x_i)) \quad (67)$$

Debido a un posible sobreajuste de f en una hipotética clase H , se añade un término de capacidad de control que regulariza el riesgo funcional:

$$R_{reg}[f] = R_{emp}[f] + \frac{\lambda}{2} \|w\|^2, \quad (68)$$

donde $\lambda > 0$ es una constante de regularización.



2.4.3.2 Máxima verosimilitud

La configuración estándar en el caso de SV es la ya descrita función de pérdida ε -insensible.

$$c(x, y, f(x)) = |y - f(x)|_\varepsilon \quad (69)$$

Nótese que minimizar (68) utilizando (69) es equivalente a minimizar (52), con la diferencia de que en este caso $C = 1/(\lambda n)$.

Para saber que función de costo utilizar se debe tener en cuenta que, por un lado, se quiere evitar utilizar un $c(x, y, f(x))$ muy complejo, ya que podría conducir a resolver problemas de optimización difíciles y, por otro lado, se debe elegir una función de costo que mejor se adapte al problema.

Bajo el supuesto de que las muestras han sido generadas por una dependencia funcional subyacente más un ruido, es decir $y_i = f_{true}(x_i) + \xi_i$ con densidad $p(\xi)$, entonces la función de costo óptima en el sentido de máxima verosimilitud es:

$$c(x, y, f(x)) = -\log p(y - f(x)) \quad (70)$$

2.4.3.3 Funciones de costo y modelos de densidad

Además de la función de pérdida ϵ -insensible, existen otras funciones comúnmente usadas como las funciones de pérdida Laplace, cuadrática y Huber.

La función de pérdida en la Figura 8 (a) corresponde al criterio convencional de error de mínimos cuadrados. La función de pérdida Laplaciana (b) es menos sensible a los *outliers* que la cuadrática. Huber propuso la función de pérdida en (c) como una función de pérdida robusta que presenta propiedades óptimas cuando se desconoce la distribución subyacente de datos. Por último, la función de pérdida ϵ -insensible (d), descrita en la sección anterior, fue propuesta por Vapnik como una aproximación a la función de pérdida de Huber, que permite obtener un conjunto escaso de vectores de soporte, a diferencia de las otras [35].

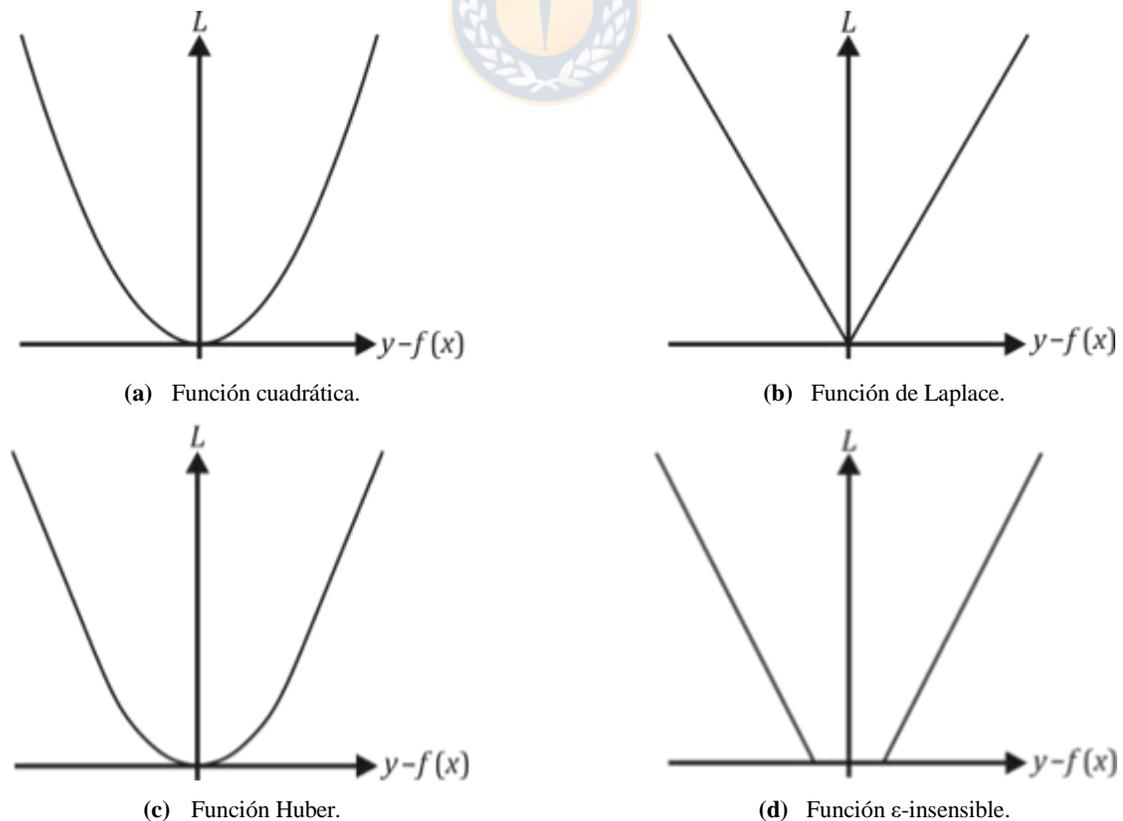


Figura 8. Funciones de pérdida.

La siguiente tabla contiene una visión general sobre algunos modelos de densidad y su correspondiente función de pérdida definida por la ecuación (70).

Tabla 2. Funciones de pérdida comunes y sus correspondientes modelos de densidad.

	Función de pérdida	Modelo de densidad
ε-insensible	$c(\xi) = \xi _\varepsilon$	$p(\xi) = \frac{1}{2(1 + \varepsilon)} \exp(- \xi _\varepsilon)$
Laplaciana	$c(\xi) = \xi $	$p(\xi) = \frac{1}{2} \exp(- \xi)$
Gaussiana	$c(\xi) = \frac{1}{2} \xi^2$	$p(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right)$
Huber robusto	$c(\xi) = \begin{cases} \frac{1}{2\sigma} (\xi)^2 & \text{si } \xi \leq \sigma \\ \xi - \frac{\sigma}{2} & \text{en otro caso} \end{cases}$	$p(\xi) \propto \begin{cases} \exp\left(-\frac{\xi^2}{2\sigma}\right) & \text{si } \xi \leq \sigma \\ \exp\left(\frac{\sigma}{2} - \xi \right) & \text{en otro caso} \end{cases}$
Polinomial	$c(\xi) = \frac{1}{p} \xi ^p$	$p(\xi) = \frac{p}{2\Gamma(1/p)} \exp(- \xi ^p)$
Polinomial a trozos	$c(\xi) = \begin{cases} \frac{1}{p\sigma^{p-1}} (\xi)^p & \text{si } \xi \leq \sigma \\ \xi - \sigma \frac{p-1}{p} & \text{en otro caso} \end{cases}$	$p(\xi) \propto \begin{cases} \exp\left(-\frac{\xi^p}{p\sigma^{p-1}}\right) & \text{si } \xi \leq \sigma \\ \exp\left(\sigma \frac{p-1}{p} - \xi \right) & \text{en otro caso} \end{cases}$

2.5 CATEGORIZACIÓN DE RECURSOS MINERALES

La categorización de recursos minerales es una de las etapas críticas a la hora de realizar un proyecto minero, ya que de esta depende el financiamiento e inversiones que se realicen en este. Es una etapa subjetiva ya que el resultado depende directamente de la persona competente⁵ encargada de la evaluación.

A lo largo de los años, se han implementado diversos códigos internacionales para categorizar recursos mineros, siendo el más destacado el código JORC (*Joint Ore Reserves Committee of The Australasian Institute of Mining and Metallurgy, Australian Institute of Geoscientists and Minerals Council of Australia*) [36].

⁵ Profesional de la industria minera que es miembro de una “Organización Profesional Reconocida”, quien debe tener un mínimo de 5 años de experiencia ejerciendo en un yacimiento con el estilo de mineralización o tipo de depósito en consideración.

Los recursos minerales pueden ser catalogados como Inferidos, Indicados y Medidos según la confianza geológica que se tenga [2, 36], donde el recurso mineral inferido tiene un nivel de confianza menor a los recursos indicados y los recursos indicados tienen mayor confianza que los recursos inferidos pero menor confianza que los recursos medidos, tal como se visualiza en la Figura 9. Los recursos minerales indicados pueden transformarse en Reservas de Mena⁶ probables, mientras que los recursos medidos pueden transformarse a reservas de mena probables o probadas, dependiendo de la decisión de la persona competente debido a la incertidumbre asociada a los factores modificadores tomados en cuenta en la conversión de recursos minerales a reservas de mena.



Figura 9. Relación entre recursos minerales y reservas de mena.

2.5.1 METODOLOGÍA DE CATEGORIZACIÓN DE RECURSOS MINERALES

El método de estimación kriging y sus variantes entregan una medida de error de estimación identificada como varianza de kriging (cokriging), donde esta puede ser utilizada como base para el establecimiento de una categoría de recurso mineral.

⁶ Parte económicamente explotable de un Recurso Mineral Indicado y/o Medido.

Diehl y David propusieron una metodología para la categorización de recursos minerales (Tabla 3) [37], definiendo niveles de incertidumbre (confianza) y precisiones (tolerancia de error) para cada caso. La precisión se basa en la desviación estándar relativa de kriging $\sigma(x_0)/Z_1^*(x_0)$ según la siguiente ecuación:

$$Precisión = \frac{\sigma(x_0) \times 100 \times \mu_{\alpha/2}}{Z_1^*(x_0)}, \tag{71}$$

donde $\sigma(x_0)$ es la desviación estándar de kriging, $Z_1^*(x_0)$ es el estimador de (co)kriging y $\mu_{\alpha/2}$ se refiere al valor de la tabla de distribución normal estándar con una cola, con una confianza de $(1 - \alpha) \times 100\%$.

Tabla 3. Clasificación de recursos minerales.

	Identificado				Sin descubrir	
	Demostrado		Indicado	Inferido	Hipotético	Especulativo
	Medido		Probable			
	Probado	Probable	Probable			
Tolerancia	± 10%	± 20%	± 40%	± 60%		
Confianza	> 80%	60 – 80 %	40 – 60 %	20 – 40%	10 – 20%	< 10%
	Recursos económicamente significativos					

CAPÍTULO III: CASO SIMULADO

3.1 METODOLOGÍA DE TRABAJO

El procedimiento de trabajo con el conjunto de datos simulados es el siguiente:

- **Creación del espacio de trabajo:** según un largo, ancho y alto establecido de un cuerpo mineralizado, se crean las coordenadas cartesianas de puntos distribuidos a lo largo de este, las que son separadas entre sí por un bloque de tamaño definido.
- **Simulación de base de datos:** para generar la base de datos, se crea un conjunto de 100 co-simulaciones gaussianas que contienen tres variables continuas, representando leyes de cobre, molibdeno y oro, donde, mediante los variogramas directos y cruzados, se define una alta correlación entre variables y una baja correlación espacial.
- **Particionamiento de la base:** para la posterior estimación y validación, las bases generadas se particionan en dos, una para utilizarla en el entrenamiento de los algoritmos y otra para la validación de estos. Para ello, se extraen una cantidad de sondajes con diferentes largos y diferente cantidad de información contenida en ellos. Los datos sobrantes son los utilizados para realizar la validación.
- **Métodos de estimación:** se utilizan cuatro algoritmos de estimación para realizar la predicción de las variables continuas en las coordenadas contenidas en la base de validación, donde se utiliza un método tradicional: Cokriging Ordinario, y tres métodos de *Machine Learning*: Redes Neuronales Artificiales, *Support Vector Regression* y Árboles de Regresión. Para cada uno de estos métodos, se debe encontrar la combinación de valores óptimos para entrenar la base y ajustar de la mejor manera los algoritmos para disminuir el error de estimación.
- **Medidas de error de estimación:** se realiza un estudio comparativo de los métodos de estimación según las medidas de error de estimación utilizadas en el trabajo de investigación (ME, MAE, RMSE y R^2).
- **Categorización de recursos minerales:** se genera una categorización de recurso mineral según la varianza de cokriging obtenida para el mineral de interés, evaluando la calidad de los datos para cada clasificación.

- **Análisis económico:** se realiza un análisis económico en las cien simulaciones, tanto para las estimaciones como para el valor real, con el objetivo de estudiar la incertidumbre de los métodos utilizados.

3.2 APLICACIÓN

3.2.1 GENERALIZACIÓN DEL ESPACIO DE TRABAJO

La primera etapa que se debe realizar es la creación del espacio de trabajo en el cual se van a realizar las simulaciones de las variables continuas. Para ello, se crea una grilla que se expande desde los puntos (0, 0, 0) hasta (100, 150, 250), avanzando de 10 en 10, lo que finalmente representa un yacimiento con dimensiones 150 [m] de ancho, 100 [m] de largo y 250 [m] de profundidad, donde el tamaño de bloque es de 10 x 10 x 10 [m³], obteniendo un total de 4,576 puntos, lo que se visualiza en la Figura 10.

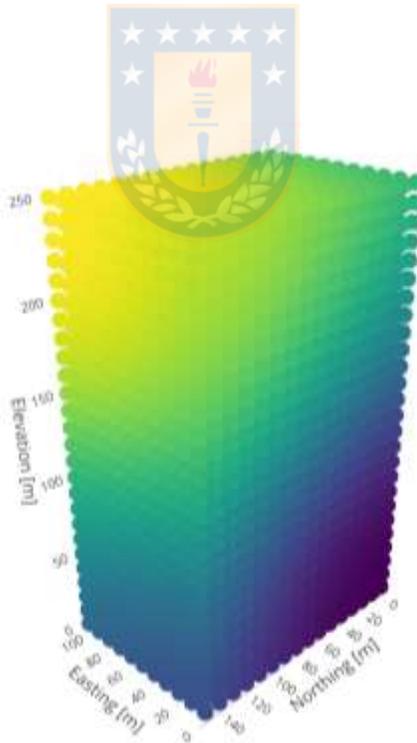


Figura 10. Espacio de trabajo.

3.2.2 SIMULACIÓN DE VARIABLES CONTINUAS

Con el fin de crear la base de datos, se hacen 100 realizaciones de co-simulaciones gaussianas de media cero en las coordenadas del yacimiento previamente creadas, donde se simulan las leyes de cobre, molibdeno y oro. El número de simulaciones es escogido con el fin de crear diferentes escenarios posibles y lograr un resultado más representativo.

La simulación de las variables continuas se realiza utilizando variogramas simples y cruzados previamente definidos, los que son creados en función de obtener una baja correlación espacial y una alta correlación entre variables. Debido a que es complejo replicar dichas condiciones, los parámetros para crear los variogramas son escogidos tratando de asemejarse lo máximo posible al escenario requerido, los que se resumen en la Tabla 4 y la visualización de los variogramas se encuentra en la Figura 11.

Tabla 4. Parámetros de variogramas simples y cruzados utilizados para crear las simulaciones de variables continuas.

	Pepa	Meseta	Rango	Modelo
Cobre	0.6	0.95	50	Gaussiano
Molibdeno	0.6	0.92	50	Gaussiano
Oro	0.6	0.91	50	Gaussiano
Cobre – Molibdeno	0.001	0.9	50	Gaussiano
Cobre – Oro	0.001	0.9	50	Gaussiano
Molibdeno – Oro	0.001	0.9	50	Gaussiano

Cabe señalar que el rango y el modelo de los variogramas es el mismo para todas sus variables con el fin de que el modelo lineal de correogionalización se cumpla y las simulaciones se puedan ejecutar. Con los valores de la tabla anterior, se crea un ciclo que solo guardan las bases de datos que contengan una correlación entre variables mayor o igual a 0.6. Una vez obtenida las cien bases de datos, se llevan las leyes simuladas a una escala normalizada en un rango de 0 a 1, según la ecuación (72), en la cual las leyes de cobre son medidas en porcentaje [%], y las leyes de oro y molibdeno son medidas en partes por millón [ppm], donde los valores de esta última son multiplicados por 100 para generar leyes de 0 a 100 [ppm]. Finalmente, la correlación obtenida entre las variables varía entre 0.6 y 0.67 aproximadamente. En la Figura 12 se muestran las leyes de cobre, molibdeno y oro de la

simulación uno distribuidos a lo largo de las coordenadas en 3D, cuyas correlaciones entre variables son de 0.615 entre el cobre y molibdeno, 0.631 entre el cobre y oro y 0.62 entre el molibdeno y el oro.

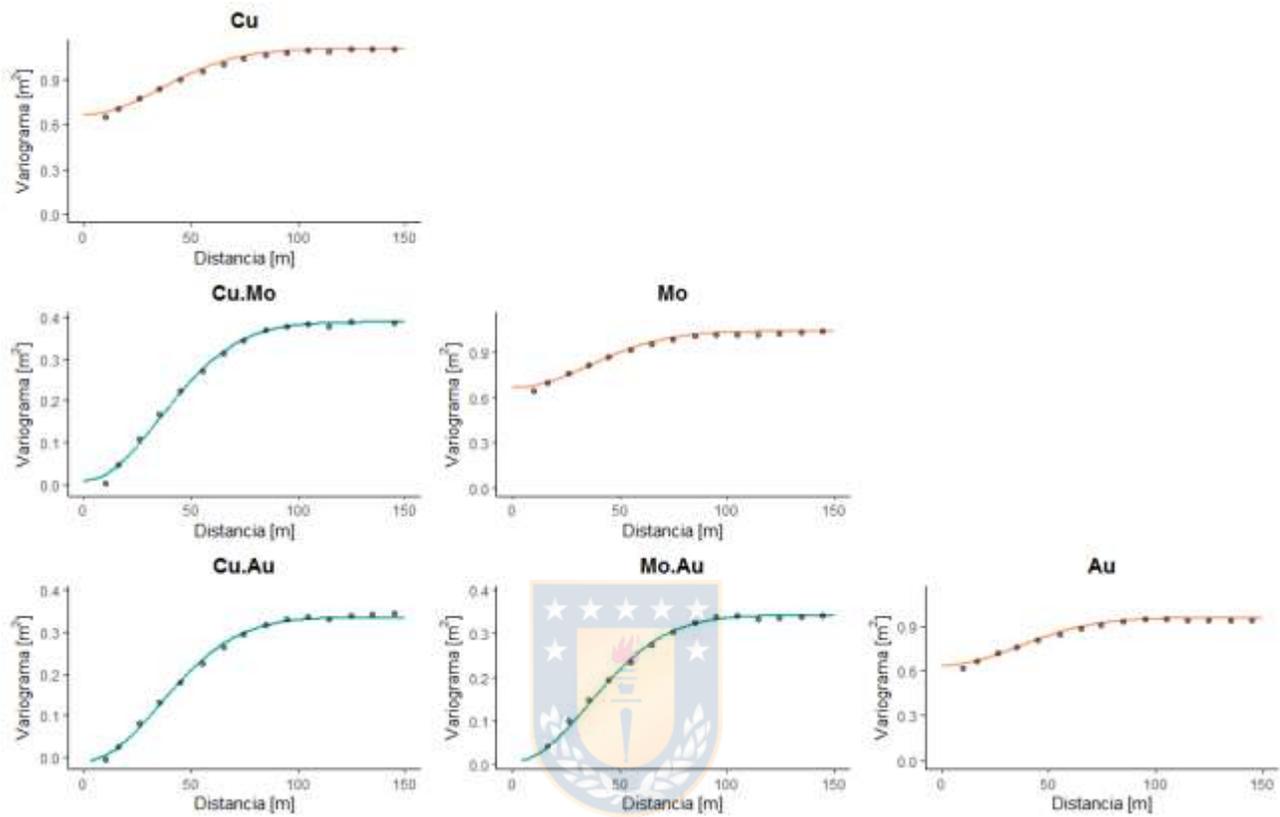
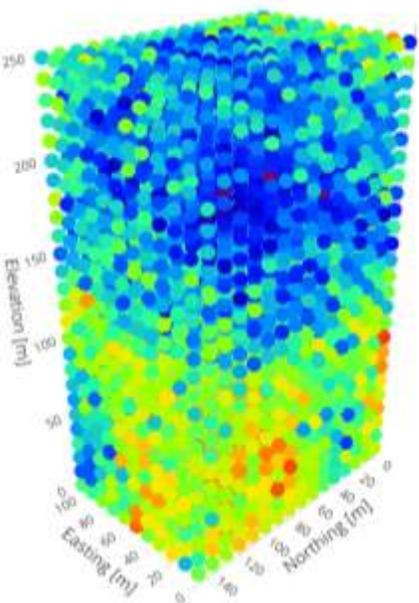
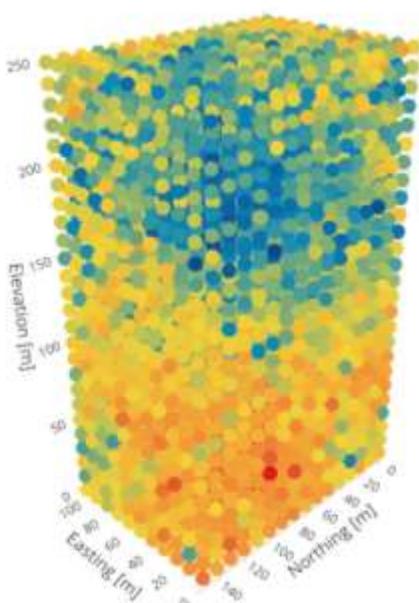


Figura 11. Variogramas simples y cruzados utilizados para hacer la simulación de las variables continuas. Puntos: variogramas experimentales. Líneas: modelos ajustados.

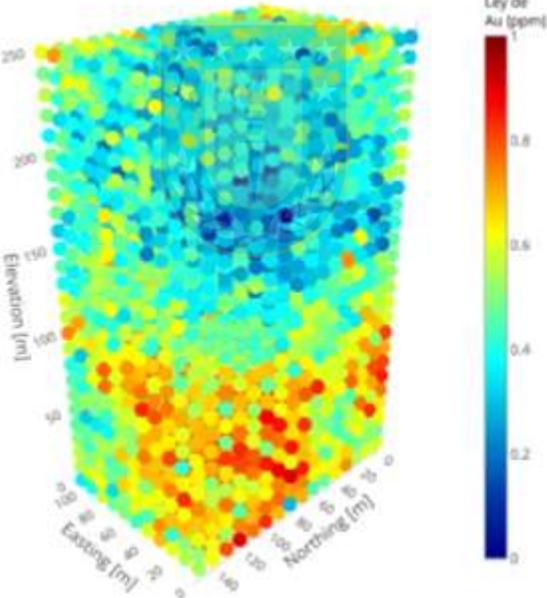
$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (72)$$



(a) Ley de cobre.



(b) Ley de molibdeno.



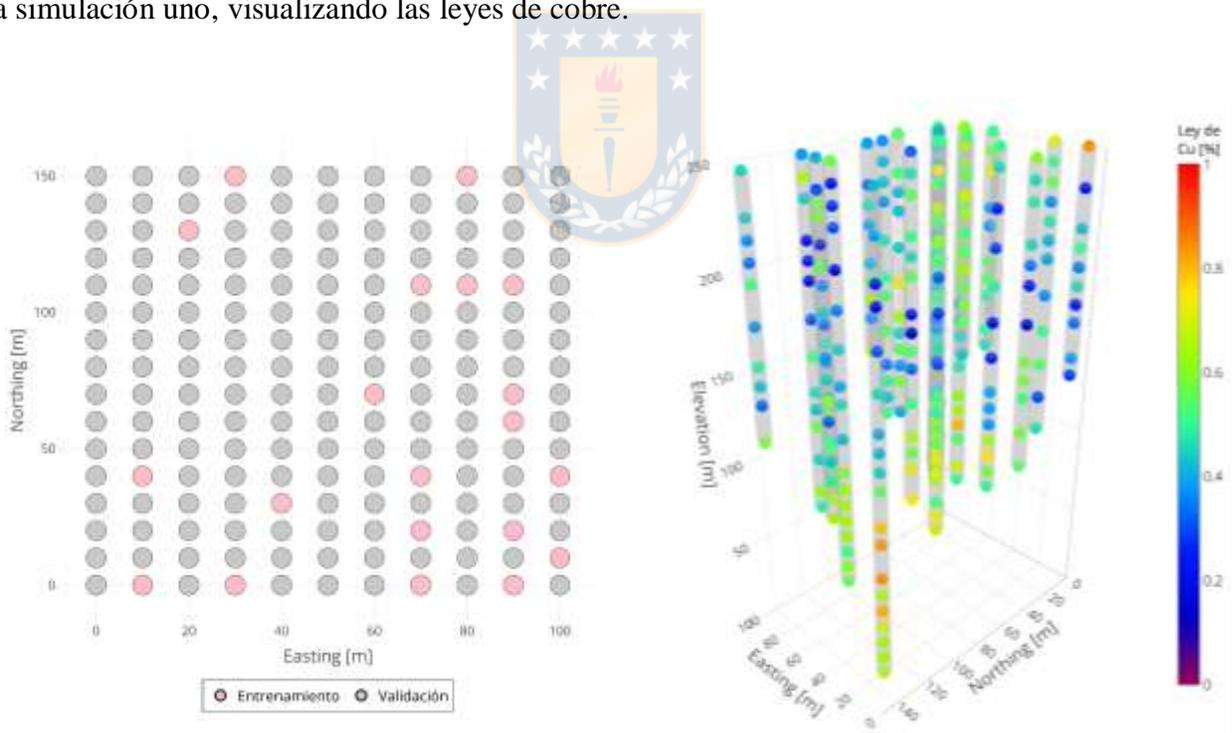
(c) Ley de oro.

Figura 12. Leyes de cobre, molibdeno y oro de la simulación uno.

3.2.3 PARTICIONAMIENTO DE BASE

La base generada requiere ser particionada en dos para obtener una base de entrenamiento, necesaria para ajustar los parámetros de los métodos de estimación, y otra de validación, utilizada para la calibración de estos.

Uno de los requerimientos de este trabajo de investigación es trabajar con pocos datos. Bajo esta premisa, se seleccionan 20 sondajes al azar, donde cada base simulada posee una distribución diferente de extracción de sondajes, cuyos largos varían de 100 a 250 [m]. Además, son extraídos algunos datos dentro de los sondajes con la intención de asemejarse a la realidad, considerando que los testigos no siempre son continuos. Finalmente, estos sondajes componen la base de entrenamiento, con una cantidad que varía entre los 247 y 312 datos, representando del 5.4 al 6.8% de la base total, y los datos sobrantes pasan a formar la base de validación. En la Figura 13 se muestran los sondajes extraídos de la simulación uno, visualizando las leyes de cobre.



(a) Sondajes desplegados en *Easting* y *Northing*.

(b) Sondajes desplegados en 3D.

Figura 13. Sondajes extraídos de la simulación uno.

Es preciso tener en cuenta que, para que los datos de entrenamiento se ajusten de la mejor manera al modelo de estimación, los conjuntos de datos generados y la base sin particionar deben ser

estadísticamente similares entre sí. Para asegurar lo anterior, se utiliza el test de Análisis de Varianza (ANOVA por sus siglas en inglés), cuya hipótesis nula de la prueba dictamina que todas las medias de los grupos son iguales mientras que la hipótesis alternativa dispone que al menos uno de los grupos tiene media diferente al resto. En la Figura 14 se muestra un gráfico de densidad luego de realizar la partición utilizando ANOVA para comparar la distribución de datos de la base total, la base de validación y la base de entrenamiento para la simulación uno, donde se puede apreciar que, para cada variable continua, las bases son similares entre sí.

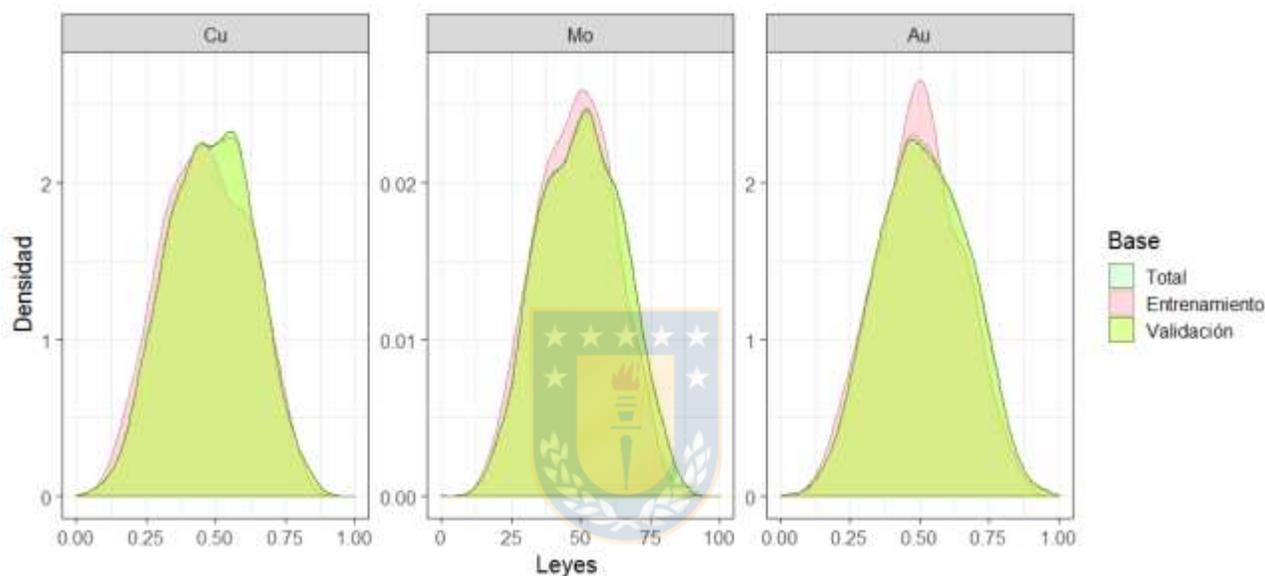


Figura 14. Gráfico de densidad para la base total, de entrenamiento y validación de la simulación uno.

3.2.4 MÉTODOS DE ESTIMACIÓN

A continuación, se describen las etapas desarrolladas para realizar la estimación de cada variable continua según el método utilizado.

3.2.4.1 Cokriging Ordinario

Para realizar la estimación con el método de estimación multivariable clásico de geoestadística Cokriging Ordinario (OCOK) es necesario en primera instancia estudiar los mapas variográficos, con el fin de determinar si existe anisotropía. En la Figura 15 se aprecian los mapas variográficos de la base de entrenamiento de la simulación uno, en el que se puede desprender que no existe cambio de

continuidad espacial por dirección, por lo tanto, no es necesario utilizar variogramas direccionales. Siendo un caso isótropo basta con utilizar variogramas omnidireccionales, los que son destacados en la Figura 16 que contiene los variogramas omnidireccionales experimentales de la base de entrenamiento de la simulación uno con su modelo ajustado. Cabe destacar que para todos los casos simulados no existe anisotropía.

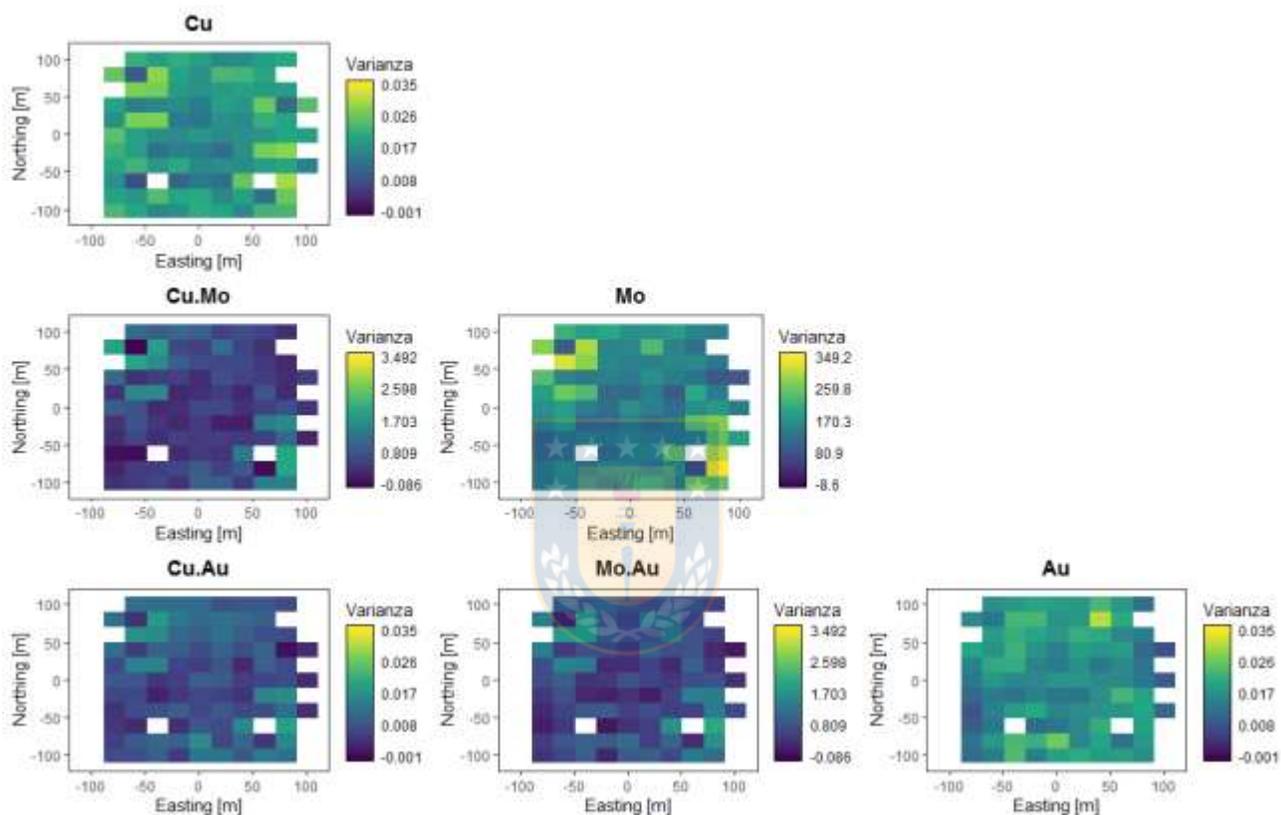


Figura 15. Mapas variográficos de la base de entrenamiento de la simulación uno.

Una vez que los variogramas simples y cruzados estén determinados y ajustados, se hace la predicción con Cokriging Ordinario en los puntos de la base de validación, donde se debe señalar el número máximo de observaciones cercanas para estimar un dato, el que puede corresponder a un solo dato como al total de los datos disponibles. Después de una serie de pruebas, se determina que el valor con el que se obtiene mejores resultados es utilizando diez vecinos cercanos. Al obtener las estimaciones se procede a calcular el error de estimación, cuyas medidas y ecuaciones se especifican en la sección 3.2.5

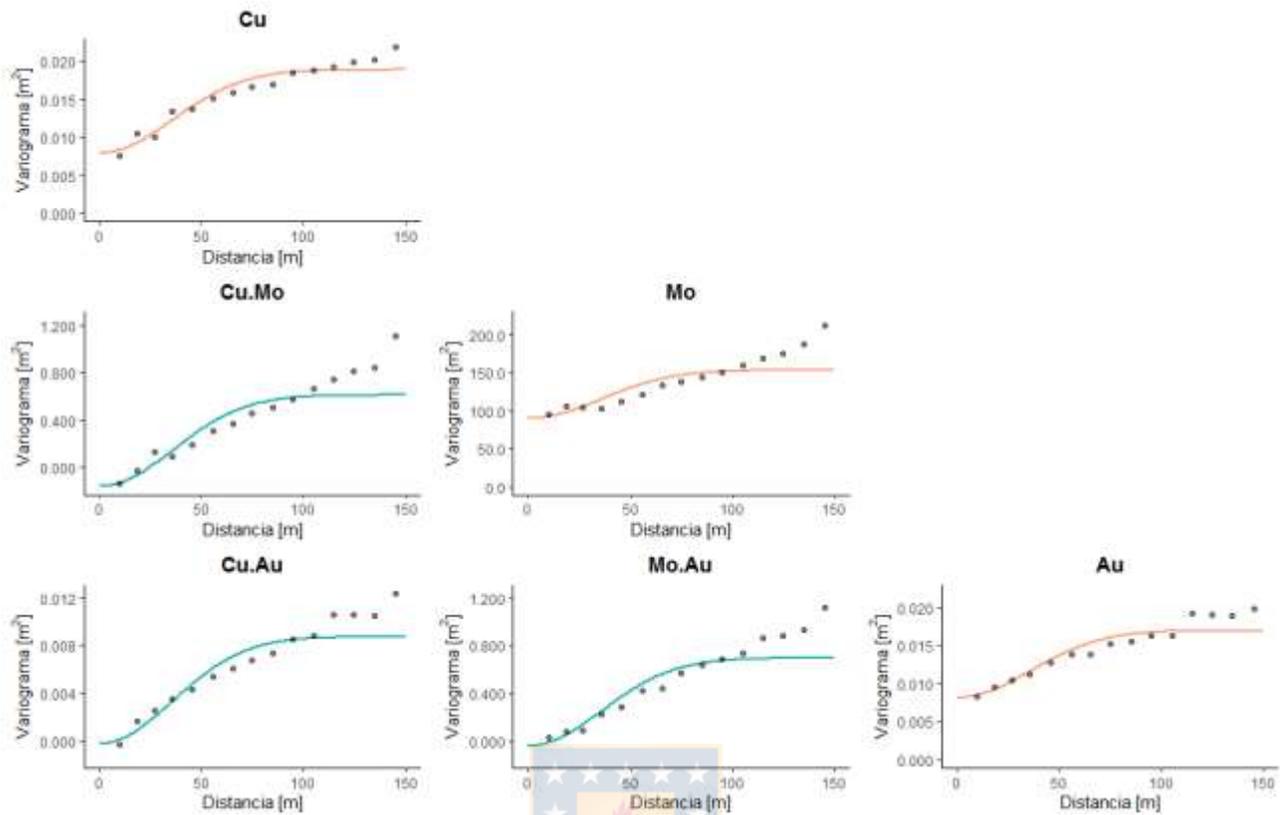
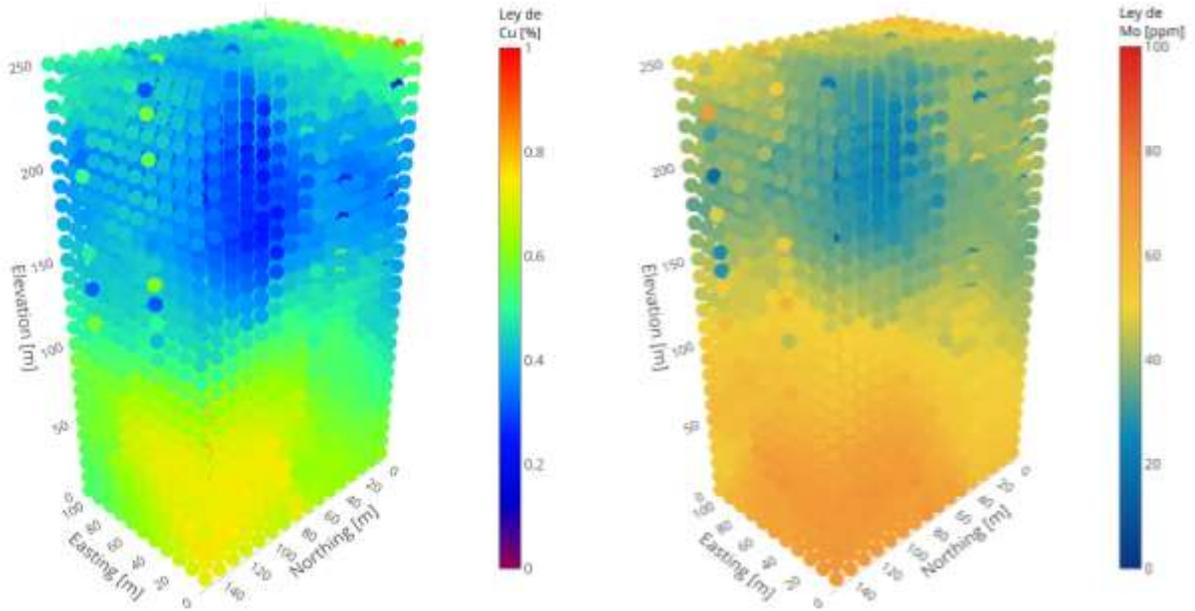


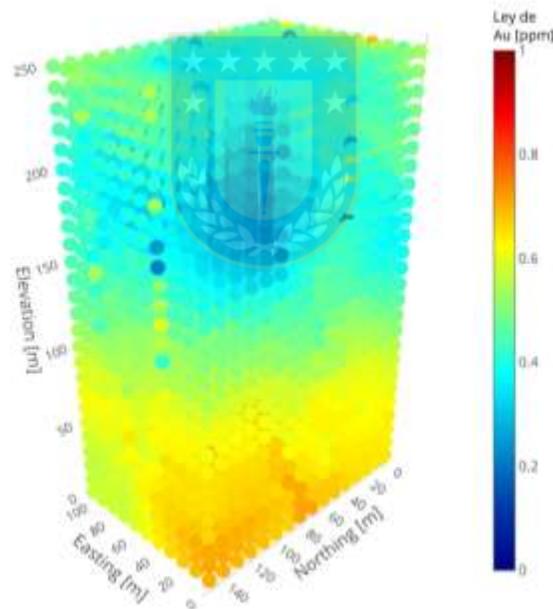
Figura 16. Variogramas omnidireccionales de la base de entrenamiento para la simulación uno. Puntos: variogramas experimentales. Líneas: modelos ajustados.

Es importante recordar que, tal como en la simulación de las variables continuas, los variogramas simples y cruzados deben modelarse con el mismo rango y modelo para que el MLC exista y la predicción sea posible, donde se utiliza un rango de 50 y el modelo Gaussiano para las cien simulaciones. Por último, la estimación con Cokriging Ordinario de las leyes de cobre, molibdeno y oro para la simulación uno se observa en la Figura 17, en la que, como primera instancia, se puede observar que las leyes en las tres variables han sido suavizadas.



(a) Ley de cobre.

(b) Ley de molibdeno.



(c) Ley de oro.

Figura 17. Estimación con Cokriging Ordinario para la simulación uno.

3.2.4.2 Redes Neuronales Artificiales

En el caso de la estimación utilizando *Machine Learning*, se deben tener ciertas consideraciones. Si se quiere estimar la ley de cobre, por ejemplo, con la información de las coordenadas y de la ley de molibdeno y/u oro, se debe tener en cuenta que en la base de validación debe existir esta información.

En la realidad, cuando se estima un cuerpo mineralizado generalmente la información disponible en los puntos de estimación no son más que las coordenadas. Teniendo esto en cuenta es que se plantea la siguiente metodología para hacer las estimaciones con ML:

1. Seleccionar la variable que tenga mejor correlación con las otras dos según la base de entrenamiento.
2. Entrenar el modelo de estimación utilizando la base de entrenamiento para la variable seleccionada anteriormente en función de las coordenadas de la base de validación.
3. Realizar estimación univariable de la variable en la base de validación.
4. Entrenar el modelo de estimación en función de las coordenadas y de la variable seleccionada en el punto 1 de la base de validación para las dos variables restantes.
5. Estimar las dos variables en la base de validación sumándole los datos de la variable estimada en el punto 3.

En resumen, se realizan dos estimaciones, la primera es univariable mientras que la segunda es multivariable, donde es recomendable que los datos en la base de entrenamiento sean normalizados antes de realizar las estimaciones para un mejor procesamiento de estos.

Para ANN se utiliza el algoritmo *Resilient Backpropagation* (Rprop), donde existen una serie de valores que pueden ser ajustados para obtener mejores resultados de estimación. La función de activación empleada para todos los casos simulados es la función sigmoideal, la que es especificada en el Anexo A. 1. Además, los parámetros que se destacan a ser ajustados en la estimación son el número de capas ocultas, el número de neuronas en cada una de estas capas y el *threshold*, valor numérico que especifica el límite de las derivadas parciales de la función de error como criterio de detención.

Además de tener como objetivo la minimización del error de estimación, también se considera el tiempo de procesamiento. Cuando la cantidad de capas ocultas y el número de neuronas en cada una es demasiado grande y, además, el *threshold* es demasiado pequeño, el tiempo de estimación aumenta considerablemente sin obtener los mejores resultados. Después de varios procesos de prueba y error para las cien bases simuladas, se determina que el número de capas ocultas que arroja un menor error de estimación es de tres.

Para determinar el número de neuronas para cada capa oculta y el *threshold* óptimo, primero se crea una grilla variando entre 5, 7, 9, 15 y 20 neuronas y, una vez determinada la mejor combinación, se busca un $threshold = \{0.04, 0.05, \dots, 0.11\}$, pues de esta manera el tiempo de procesamiento no es tan grande. Con lo anterior, se hacen poco más de 120 combinaciones, donde se calcula el error de estimación para cada una en las cien simulaciones comparando el valor estimado con el valor real. Finalmente, los parámetros utilizados para la estimación definitiva se eligen según el menor error de estimación obtenido.

Continuando con el ejemplo de la simulación uno, la variable que tiene mejor correlación con las otras dos en la base de entrenamiento es el oro (ver Tabla 5), ya que la suma de todas sus correlaciones es mayor, por lo que es este el que se estima primero con ANN siguiendo la metodología descrita. Una vez optimizados el número de neuronas para cada capa oculta y el *threshold*, se realiza la estimación de esta. Posteriormente, se establece que el cobre y molibdeno dependen de las coordenadas y del oro, y en base a eso se hace la optimización de los parámetros para cada uno, finalizando en su estimación. Los parámetros óptimos para las estimaciones de las tres variables se encuentran en la Tabla 6 y, a modo explicativo, la estructura de las redes neuronales de estas se puede visualizar en la Figura 18.

Tabla 5. Correlación entre el cobre, molibdeno y oro en la base de entrenamiento para la simulación uno.

	Cobre	Molibdeno	Oro
Cobre	1	0.409	0.515
Molibdeno	0.409	1	0.474
Oro	0.515	0.474	1
Suma	1.924	1.883	1.989

Tabla 6. Parámetros óptimos de ANN para la simulación uno.

	Cobre	Molibdeno	Oro
N° neuronas capa 1	15	15	15
N° neuronas capa 2	5	9	15
N° neuronas capa 3	9	15	5
Threshold	0.09	0.08	0.11

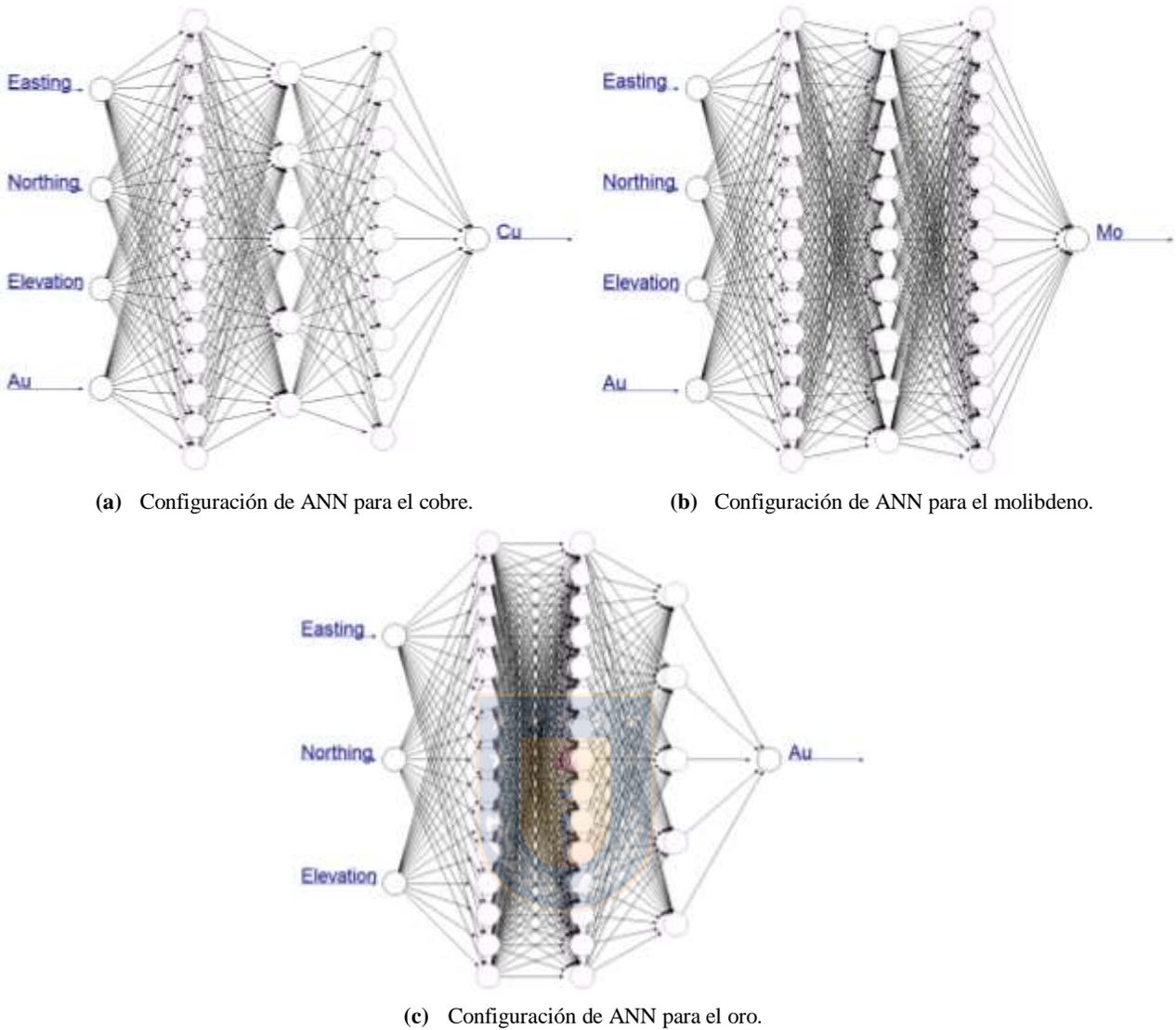
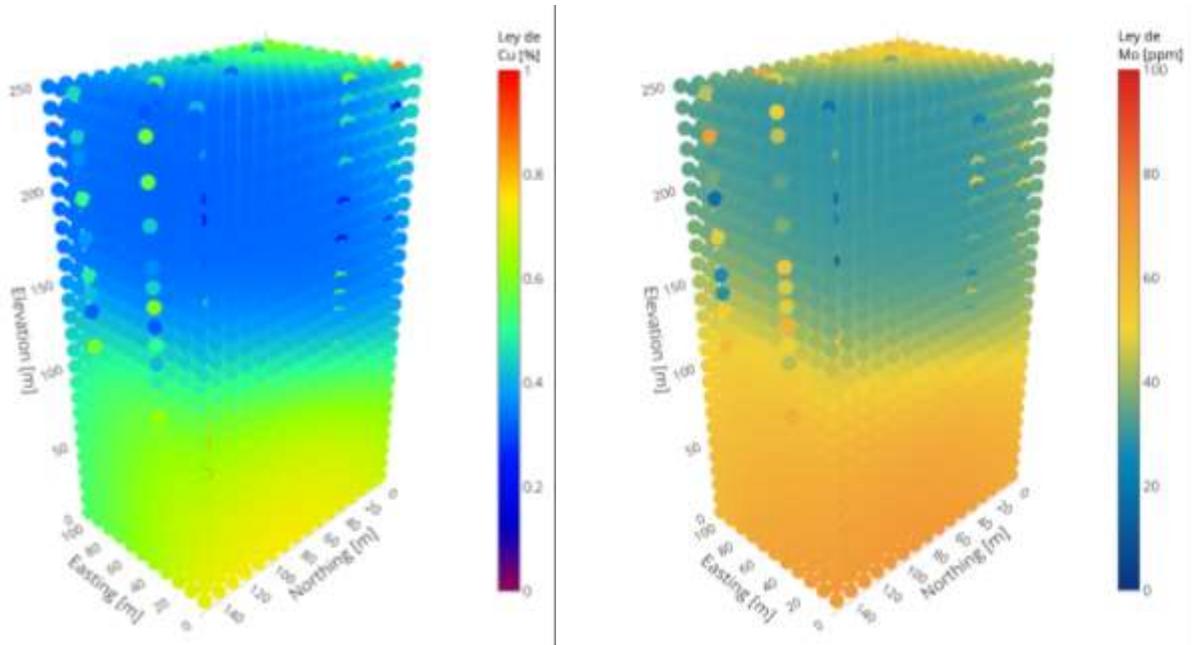


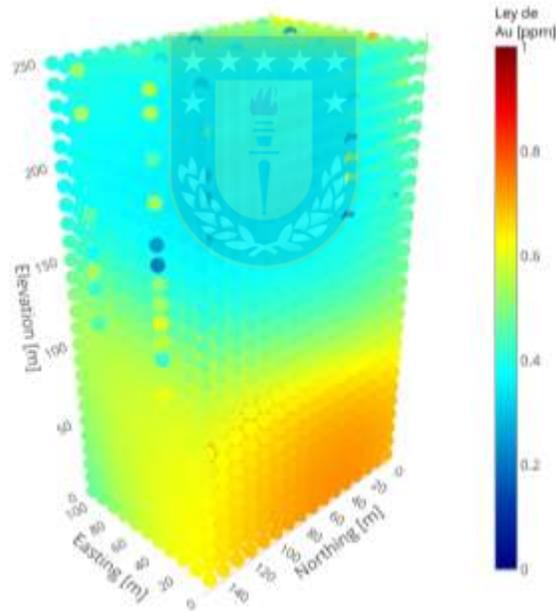
Figura 18. Configuración de ANN para la simulación uno.

Por último, la Figura 19 contiene las estimaciones hechas con ANN para las tres variables continuas, donde visualmente es posible notar que, en comparación con la estimación del Cokriging Ordinario, las leyes alcanzan valores más altos y bajos y, a su vez, tienen menos variaciones obteniendo colores más parejos.



(a) Ley de cobre.

(b) Ley de molibdeno.



(c) Ley de oro.

Figura 19. Estimación con Redes Neuronales Artificiales para la simulación uno.

3.2.4.3 Árboles de Regresión

Para encontrar los valores óptimos de las estimaciones realizadas para cada método, se realizan diversas pruebas cambiando los parámetros con el fin de encontrar su valor óptimo. Es así como se determina que las estimaciones utilizando Árboles de Regresión arroja valores de leyes en un rango

muy acotado, por lo que hacer la primera estimación univariable utilizando este método arrastraría errores muy grandes a las siguientes dos estimaciones que dependen de ella. Sabiendo esto, para disminuir los errores de estimación en este caso es que se realiza la estimación univariable utilizando Kriging Ordinario.

Tal como en el caso de Cokriging Ordinario, para hacer la estimación con Kriging se debe estudiar el mapa variográfico de la base de entrenamiento de la variable a predecir para tener conocimiento de la existencia de anisotropía, y posteriormente modelar el variograma simple de esta. En este caso, también se consideran diez vecinos cercanos para hacer la estimación de un punto para los cien casos. Para la simulación uno, según la Tabla 5, la variable que tiene mejor correlación con el resto es el oro, donde su mapa variográfico y variograma simple se encuentran en la Figura 15 y Figura 16, respectivamente. Después de determinar el modelo que mejor se ajusta al variograma experimental, se realiza su estimación utilizando Kriging Ordinario.

Luego, para realizar la estimación de las variables restantes se utiliza el método CART, donde los parámetros que destacan a ser optimizados son el número de observaciones mínimas que deben existir en un nodo para que se intente una división (*minsplit*), el número mínimo de observaciones para un nodo final (*minbucket*), la profundidad máxima de cualquier nodo del árbol final (*maxdepth*), y, por último, el parámetro de costo – complejidad (*cp*) cuya función es ahorrar tiempo de cómputo, evitando que se realicen divisiones innecesarias.

Para buscar los parámetros óptimos, se crea una grilla donde el *minsplit* y *minbucket* van de 1 a 8, la profundidad máxima va de 2 a 30 y el costo – complejidad puede alcanzar los valores 0.0005, 0.005, 0.01, 0.03, 0.05, 0.07, 0.09 y 0.1, obteniendo alrededor de 7700 combinaciones. Una vez calculado el error de estimación para las cien simulaciones, se eligen los parámetros óptimos de cada una en base al menor error obtenido. En el caso de la simulación uno, los parámetros óptimos obtenidos para las estimaciones del cobre y molibdeno se presentan en la Tabla 7, donde en la Figura 20 se visualiza el árbol final de regresión para ambas variables. De esta imagen, es posible apreciar que los parámetros óptimos en este caso arrojaron un árbol de dimensiones bastante pequeñas para ambas variables, donde las variables con mayor importancia son *Elevation* y el oro.

Tabla 7. Parámetros óptimos de CART para la simulación uno.

	Cobre	Molibdeno
<i>Minsplit</i>	1	1
<i>Minbucket</i>	1	1
<i>Maxdepth</i>	2	2
<i>Cp</i>	0.03	0.0005

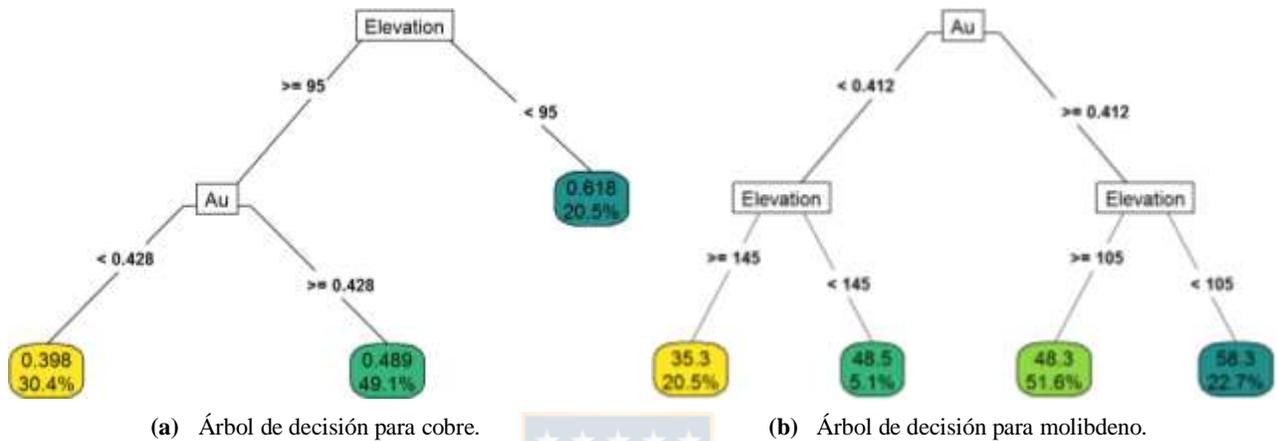
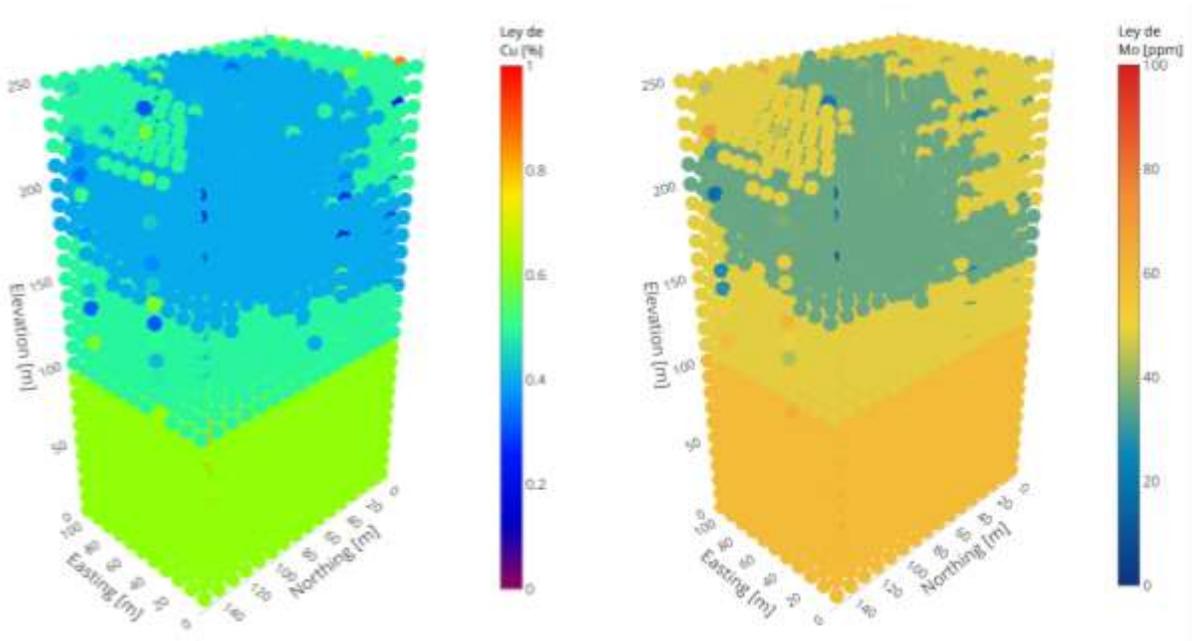


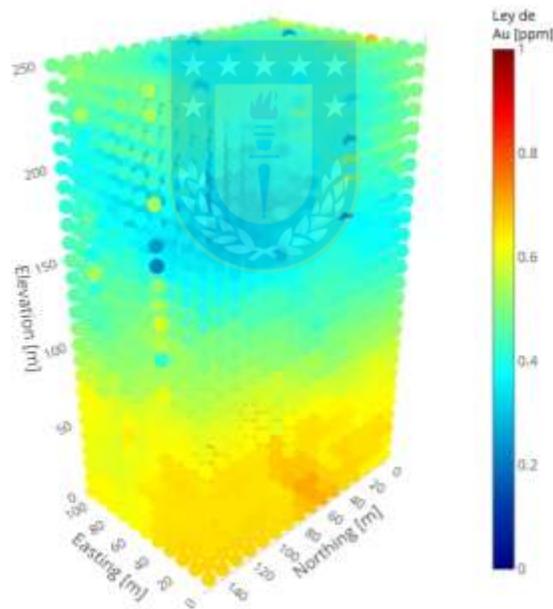
Figura 20. Árbol de decisión para la simulación uno.

Finalmente, la estimación de la simulación uno se puede apreciar en la Figura 21, recordando que el oro se estima con Kriging Ordinario de manera univariable, y el cobre y molibdeno se estiman con el método CART de forma multivariable. De la imagen se puede observar cómo los cambios en las leyes, tanto del cobre como del molibdeno, son bruscos y no continuos. A simple vista, se puede entender que estimar con este método produce errores de estimación más grande que los otros dos ya analizados.



(a) Ley de cobre.

(b) Ley de molibdeno.



(c) Ley de oro.

Figura 21. Estimación con Árboles de Regresión para la simulación uno.

3.2.4.4 Support Vector Regression

A diferencia del método CART y al igual que el método ANN, la primera variable seleccionada para ser estimada en base a las coordenadas según una mejor correlación con el resto, es estimada con

el mismo método que se utiliza para estimar las otras dos variables, que en este caso es el *Support Vector Regression*.

Luego de hacer varias pruebas de estimación, se llega a la conclusión de que el kernel que emite mejores resultados es el Gaussiano (ver Tabla 1), reescribiéndose según la siguiente ecuación:

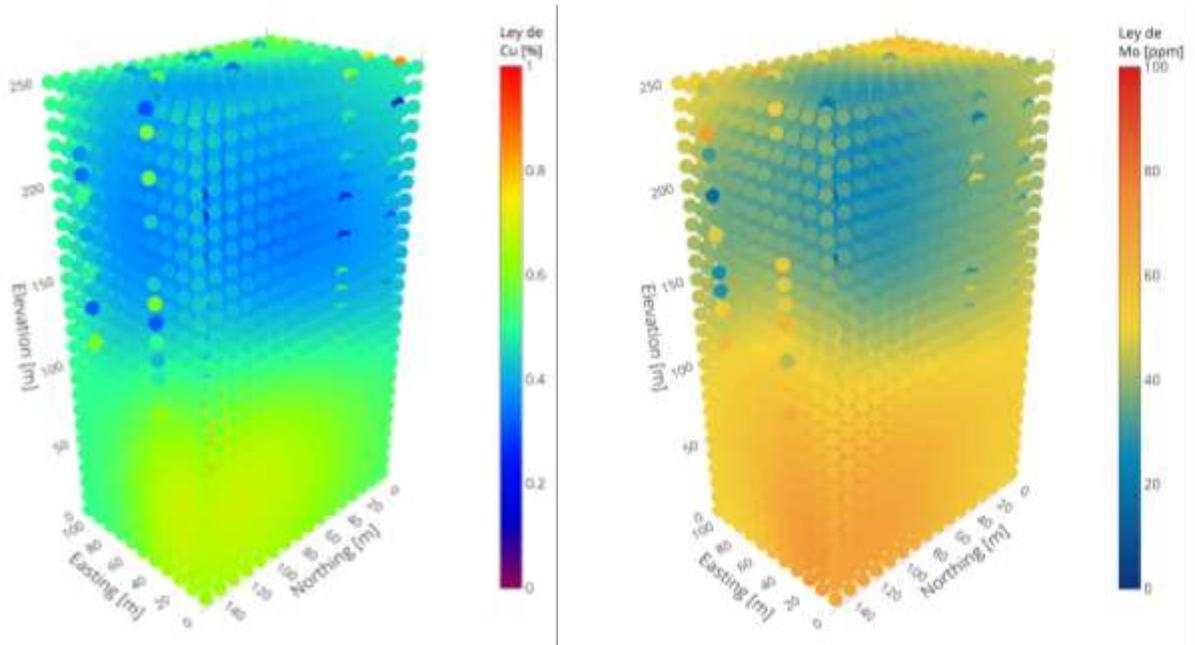
$$k(x, x') = \exp(-\gamma \|x - x'\|^2) , \quad \gamma > 0 \tag{73}$$

Los parámetros que destacan a ser optimizados para SVR son el parámetro de regularización C , épsilon (ϵ) de la función ϵ -insensible, gamma (γ) del kernel Gaussiano y la tolerancia (tol) como criterio de detención. En este caso, primero se crea una grilla con un $\epsilon = \{0.001, 0.01, 0.1, 0.2, \dots, 1.4\}$, $C = \{0.1, 0.2, \dots, 1.8, 2^1, 2^2, \dots, 2^4\}$, $\gamma = \{0.1, 0.2, \dots, 1.1\}$ y, una vez obtenida la mejor combinación según un menor error en la estimación, se busca una tolerancia que puede alcanzar los valores $\{0.07, 0.09, 0.10, 0.11, 0.1, 0.7, 0.9, 1.0, 1.1, 1.3\}$, realizando en total más de 7000 combinaciones. Al igual que ANN, se hace la optimización en dos partes pues de lo contrario el tiempo de procesamiento aumenta considerablemente. Al tener todas las combinaciones estimadas, se selecciona la que tiene menor error de estimación.

Continuando con el ejemplo de la simulación uno y como ya se ha mencionado, la primera variable a estimar es el oro en función de las coordenadas y el cobre y molibdeno se estiman en función de las coordenadas y del oro. Los parámetros de SVR con menor error de estimación se resumen en la Tabla 8 y la estimación de las tres variables se visualiza en la Figura 22, donde es apreciable que, en comparación con los otros métodos, las leyes son más suavizadas repercutiendo en menor cantidad de leyes altas, además de que los cambios entre ellas son más tenues.

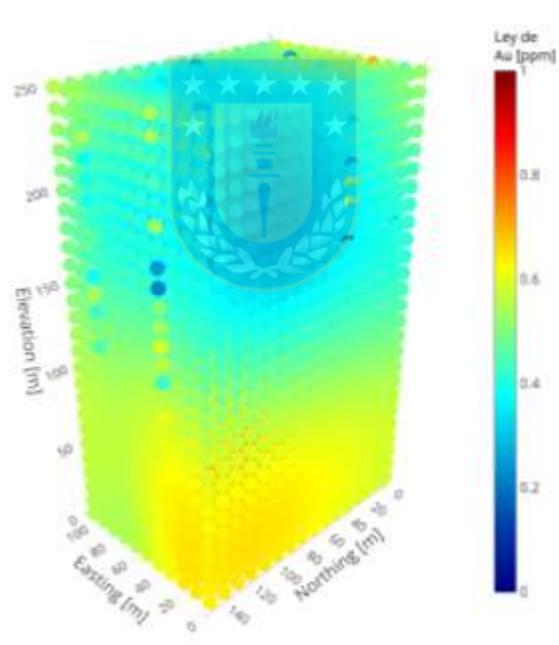
Tabla 8. Parámetros óptimos de SVR para la simulación uno.

	Cobre	Molibdeno	Oro
C	1.23	0.79	2
ϵ	1.11	0.49	0.51
γ	0.3	0.49	0.39
Tol	0.7	0.7	0.13



(a) Ley de cobre.

(b) Ley de molibdeno.



(c) Ley de oro.

Figura 22. Estimación con *Support Vector Regression* para la simulación uno.

3.2.5 MEDIDAS DE ERROR DE ESTIMACIÓN

En la sección anterior, se menciona en reiteradas ocasiones el llamado error de estimación, donde las medidas de rendimiento de los métodos consideradas en este trabajo de investigación son

el error medio (ME), el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2), cuyas ecuaciones se especifican a continuación:

$$ME = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i) \quad (74)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (75)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (76)$$

$$R^2 = \rho_{y_i, \hat{y}_i}^2, \quad (77)$$

donde y_i es la variable original, \hat{y}_i es la variable predicha y ρ es el coeficiente de correlación de Pearson. Teniendo en cuenta estas definiciones, es posible entender que se obtienen mejores resultados cuando los errores (ME, MAE y RMSE) son bajos y el R^2 es alto. Se hace un resumen estadístico de estos criterios con el fin de simplificar la comparación entre los métodos de estimación el cual se denomina *skill value*, donde se considera el ME, MAE y RMSE por igual y al R^2 se le aplica una escala para que todas las medidas tengan el mismo impacto. En la ecuación (78) se describe el *skill value* para el cobre y oro y en la ecuación (79) para el molibdeno [8, 38], donde es posible reparar en que un menor valor de esta medida se traduce en mejores resultados.

$$skill\ value_{Cu,Au} = |ME| + MAE + RMSE + 0.1 \times (1 - R^2) \quad (78)$$

$$skill\ value_{Mo} = |ME| + MAE + RMSE + 10 \times (1 - R^2) \quad (79)$$

Los *skill values* calculados para las tres variables continuas en las cien simulaciones son representados en los *boxplots* de la Figura 23, cuyos valores se encuentran en el Anexo D. 1, permitiendo visualizar un comportamiento similar para el cobre, molibdeno y oro, donde se obtienen valores menores de *skill value* para el método SVR, seguido por el método ANN y OCOK, y los peores resultados son obtenidos por el método CART.

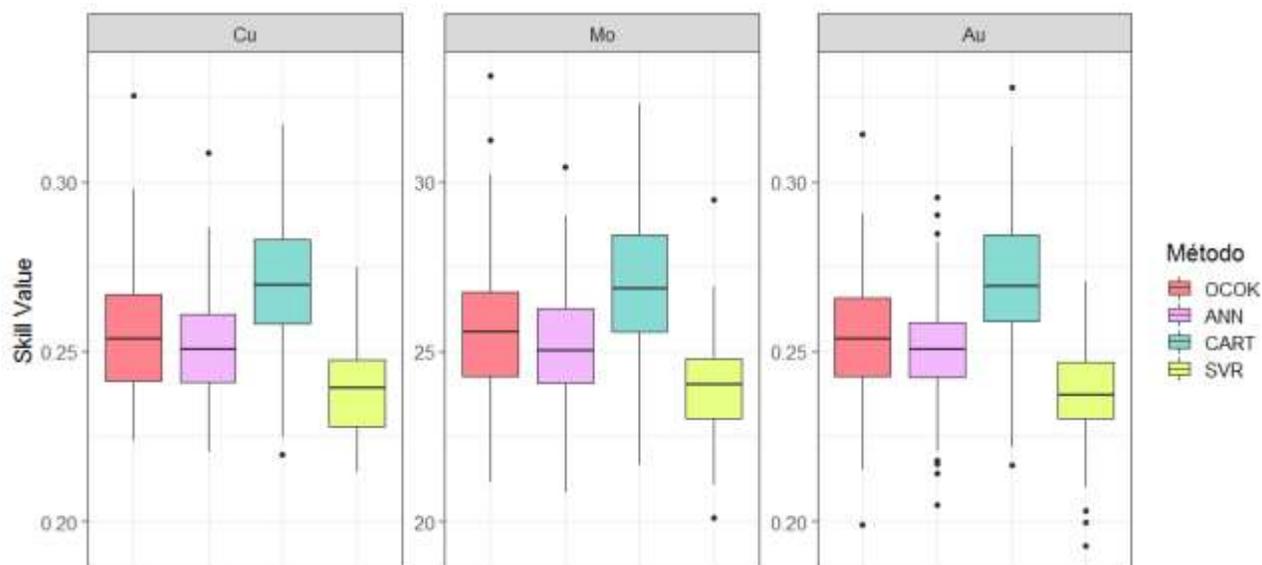


Figura 23. Gráfico de caja de los *skill value* obtenidos para cada método en las cien simulaciones.

Para el mejor, peor y más probable caso, elegidos en función de la evaluación económica cuyo procedimiento se especifica en la sección 3.2.7, los errores y coeficientes de determinación obtenidos se resumen en la Tabla 9, donde los errores del cobre y molibdeno, así como el *skill value*, están expresados en una notación científica de $\times 10^{-2}$. Se puede observar que, para los tres casos presentes en la tabla, los errores del SVR son menores que los obtenidos por los otros métodos de estimación y los coeficientes de determinación son mayores, dando como resultado *skill values* menores (mejores) y, por el contrario, el método CART es el que obtiene errores mayores y R^2 menores para el mejor y más probable caso, causando un *skill value* mayor (peor). Para el peor caso, tanto para el cobre como para el molibdeno, los peores resultados son obtenidos por el Cokriging Ordinario y, en el caso del oro, el *skill value* mayor es obtenido por el método CART.

De manera preliminar es posible mencionar que el método *Support Vector Regression* obtiene valores menores en cuanto al *skill value*, sin embargo, no es suficiente para concluir que este es mejor método que los otros evaluados. Es necesario comparar la calidad de los datos y realizar el análisis económico para estudiar la incertidumbre, resultados que se discuten en la sección 3.3.

Tabla 9. Medidas de rendimiento en las estimaciones para el mejor, peor y más probable caso.

	Mejor caso				Más probable caso				Peor caso				
	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	
Cobre	ME	0.609	1.166	1.074	0.228	1.680	1.729	1.809	0.653	4.324	3.311	3.295	1.322
	MAE	8.229	8.107	8.693	7.984	9.343	9.034	9.310	8.770	9.939	9.890	9.918	8.593
	RMSE	10.33	10.18	11.89	10.02	11.85	11.37	11.77	11.06	12.72	12.43	12.49	10.76
	R ²	0.533	0.552	0.486	0.561	0.442	0.488	0.451	0.517	0.445	0.480	0.402	0.524
	<i>Skill value</i>	23.84	23.94	25.80	22.62	28.45	27.25	28.38	25.31	32.53	30.83	31.69	25.44
	Ranking	2^{do}	3^{er}	4^{to}	1^{er}	4^{to}	2^{do}	3^{er}	1^{er}	4^{to}	2^{do}	3^{er}	1^{er}
Molibdeno	ME	0.661	-0.11	0.928	0.433	1.645	-1.97	3.082	0.774	3.106	2.001	3.221	-0.53
	MAE	7.814	7.758	8.628	7.550	9.510	9.059	10.09	8.550	10.09	9.640	10.03	8.715
	RMSE	9.851	9.786	10.86	9.502	11.98	11.43	12.78	10.72	12.74	12.19	12.65	10.88
	R ²	0.552	0.559	0.463	0.584	0.425	0.484	0.390	0.532	0.471	0.484	0.480	0.566
	<i>Skill value</i>	22.81	22.07	25.78	21.65	28.88	27.62	32.05	24.72	31.24	28.99	31.10	24.46
	Ranking	3^{er}	2^{do}	4^{to}	1^{er}	3^{er}	2^{do}	4^{to}	1^{er}	4^{to}	2^{do}	3^{er}	1^{er}
Oro	ME	0.349	-0.42	0.227	-0.28	1.316	-1.05	1.652	-0.25	3.371	3.004	3.719	0.687
	MAE	7.377	7.356	7.456	7.037	8.750	8.514	9.247	8.278	9.920	9.554	10.36	8.634
	RMSE	9.265	9.240	9.339	8.824	10.98	10.69	11.69	10.36	12.69	12.01	13.06	10.85
	R ²	0.547	0.558	0.538	0.581	0.498	0.529	0.453	0.548	0.457	0.504	0.436	0.550
	<i>Skill value</i>	21.52	21.43	21.64	20.32	26.06	24.97	28.06	23.41	31.41	29.52	32.78	24.67
	Ranking	3^{er}	2^{do}	4^{to}	1^{er}	3^{er}	2^{do}	4^{to}	1^{er}	3^{er}	2^{do}	4^{to}	1^{er}

3.2.6 CATEGORIZACIÓN DE RECURSOS MINERALES

A diferencia del kriging y sus variantes, que cuentan con una medida de error para cada bloque estimado llamada varianza de kriging, los métodos de ML no entregan una medida de riesgo de manera simultánea junto al valor estimado, por lo que se dificulta el proceso de categorización de recursos minerales. La estrategia utilizada en este trabajo de investigación es realizar la clasificación utilizando la varianza de cokriging según la metodología propuesta en la sección 2.5.1 y replicar los recursos minerales medidos, indicados e inferidos obtenidos en las coordenadas de los métodos de ML para hacer una comparación entre ellos, con el objetivo de determinar la calidad de la categorización.

Según la Tabla 3 y empleando la ecuación (71), un recurso se clasifica como medido si la desviación estándar relativa de cokriging ($\sigma(x_0)/Z_1^*(x_0)$) es menor a 0.2376 con una tolerancia de error del $\pm 20\%$ y una confianza del 60%, inferido si $\sigma(x_0)/Z_1^*(x_0)$ es mayor a 0.7628 con una

tolerancia de error del $\pm 40\%$ y una confianza del 40%, e indicado cuando $\sigma(x_0)/Z_1^*(x_0)$ es mayor a 0.2376 y menor a 0.7628. Además, se utiliza la categorización de recursos obtenida en el caso del cobre para las otras dos variables continuas, molibdeno y oro, ya que se asume que este es la mineral de interés.

Se obtuvieron recursos medidos e indicados en todas las simulaciones, sin embargo, solo en cinco simulaciones se clasificaron recursos como inferidos, con menos de 7 datos en cada una, por lo que hablar sobre la calidad de esta última categoría no resulta representativo ni concluyente. Cabe mencionar que la casi nula clasificación de recursos como inferidos se debe a lo cercano que están los sondeos al resto de los datos, cubriendo de buena manera la zona de trabajo.

La categorización resultante para el mejor, peor y más probable caso en función del porcentaje perteneciente a la base total se visualiza en la Figura 24, donde se obtiene sobre un 60% de recursos medidos y bajo un 40% de recursos indicados para el mejor y más probable caso, a diferencia del peor caso que se obtienen en cantidades similares. Además, es notable la ausencia de recursos inferidos en los tres escenarios.

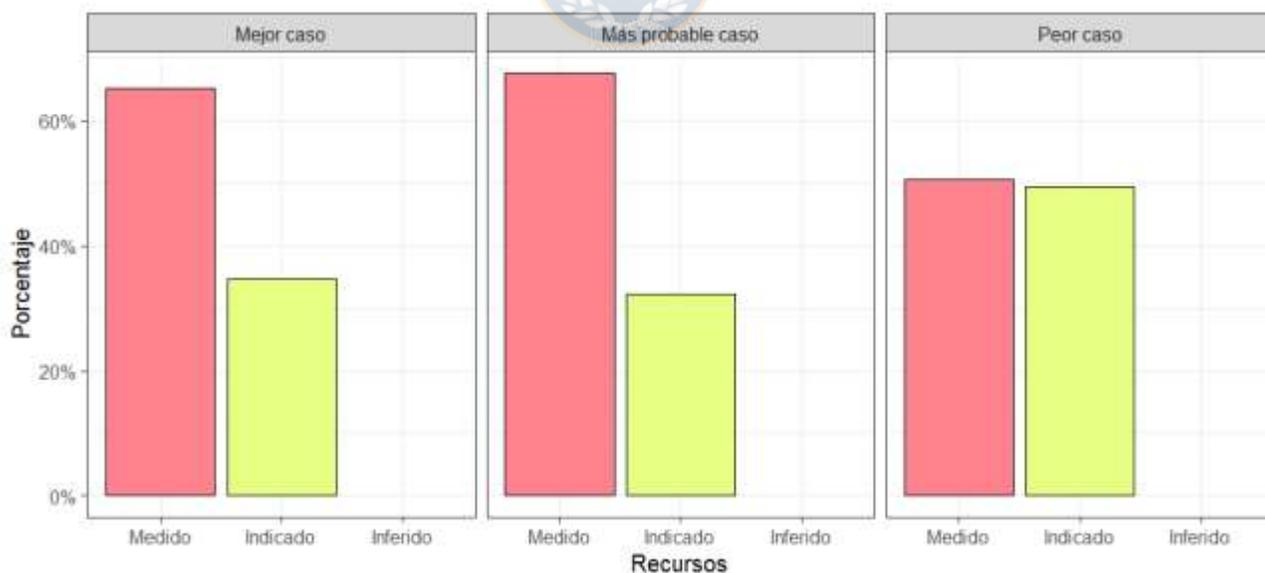


Figura 24. Categorización de recursos minerales para la estimación del mejor, peor y más probable caso.

Para determinar la calidad de los datos según el método de estimación utilizado, las principales medidas estadísticas a utilizar son la media y la desviación estándar (σ) de las leyes pues, si bien

pueden ser utilizadas muchas más medidas, producto de la complejidad de análisis solo se consideran las mencionadas. Se hace una comparación entre estas medidas de la base original y las bases estimadas con los diferentes métodos empleados y el criterio a utilizar para determinar cuál método es más o menos conservador es el siguiente:

- Mientras menor sea la ley media de la base estimada en comparación con la ley media de la base real y, además, menor sea la desviación estándar de las leyes, se dice que el método es más conservador.
- Mientras más parecida o mayor sea la ley media estimada en comparación con la ley media de la base real y, además, mayor sea la desviación estándar de las leyes, se dice que el método es menos conservador.

Las medidas estadísticas consideradas en el mejor, peor y más probable caso para determinar la calidad de los datos se resumen en la Tabla 10. En general, se puede observar que las leyes medias de los recursos medidos son mayores a las leyes medias de los recursos indicados y, por otro lado, las desviaciones estándar en la base original son notoriamente mayores que en el resto de las bases estimadas, esto debido al suavizamiento que se produce al realizar la estimación en cada método. Además, para los recursos medidos en los tres casos estudiados, se puede apreciar cierta propensión a leyes medias menores y desviaciones estándar mayores para el método CART, leyes medias mayores para ANN y desviaciones estándar menores para el método SVR. Por otro lado, el comportamiento de los datos para los recursos indicados es diferente, pues en los tres casos se obtienen leyes medias mayores con SVR, mientras que las leyes medias menores son obtenidas con OCOK en el mejor caso y con CART en el peor y más probable. En cuanto a las desviaciones estándar para los recursos indicados, existe una tendencia a desviaciones menores utilizando mayormente OCOK y desviaciones mayores utilizando SVR y ANN.

Siguiendo el criterio descrito anteriormente para determinar la calidad de los datos según el método de estimación, se establece un sistema de evaluación para determinar cuál método es más o menos conservador, obteniendo el 1^{er} lugar el método más conservador y 4^{to} lugar el menos conservador. En la Tabla 11 se encuentra el ranking obtenido para el mejor, peor y más probable caso, donde se puede observar que para los recursos medidos, en los tres casos, el método más conservador es SVR y el menos conservador es ANN. Por otro lado, para los recursos indicados, en el mejor caso

el método más conservador es OCOK, mientras que el menos conservador es SVR, siendo esto diferente para el peor y más probable caso, ya que el método más conservado es CART, en tanto que el menos conservador nuevamente es SVR.

Tabla 10. Calidad de los datos para la estimación del mejor, peor y más probable caso según recursos minerales.

			Medidos					Indicados				
			Original	OCOK	ANN	CART	SVR	Original	OCOK	ANN	CART	SVR
Mejor caso	Cu	Media	0.570	0.565	0.563	0.555	0.564	0.383	0.376	0.364	0.381	0.389
		σ	0.120	0.073	0.069	0.080	0.067	0.125	0.056	0.072	0.070	0.062
	Mo	Media	61.67	61.04	61.58	60.19	60.45	43.20	42.48	43.67	43.29	44.24
		σ	11.67	7.101	7.556	7.914	6.789	11.80	7.698	9.692	8.205	6.944
	Au	Media	0.600	0.603	0.610	0.603	0.603	0.428	0.411	0.420	0.415	0.429
		σ	0.109	0.069	0.079	0.067	0.068	0.110	0.066	0.077	0.068	0.066
Más probable caso	Cu	Media	0.633	0.611	0.614	0.611	0.614	0.445	0.438	0.430	0.435	0.464
		σ	0.120	0.066	0.059	0.069	0.058	0.150	0.049	0.071	0.046	0.082
	Mo	Media	56.88	55.27	59.51	54.29	55.69	38.32	36.59	38.87	34.19	38.42
		σ	11.87	7.391	5.398	7.630	6.336	15.03	5.272	9.523	7.376	8.854
	Au	Media	0.617	0.603	0.635	0.609	0.615	0.429	0.418	0.424	0.396	0.442
		σ	0.118	0.066	0.068	0.074	0.062	0.142	0.062	0.089	0.072	0.093
Peor caso	Cu	Media	0.640	0.627	0.644	0.612	0.620	0.466	0.391	0.395	0.428	0.459
		σ	0.126	0.082	0.102	0.090	0.081	0.130	0.044	0.063	0.017	0.066
	Mo	Media	64.25	63.87	64.21	63.90	64.38	45.58	39.68	41.57	39.42	46.52
		σ	13.00	9.016	8.720	9.009	8.499	14.19	5.703	5.944	5.331	8.861
	Au	Media	0.617	0.609	0.611	0.603	0.608	0.434	0.374	0.379	0.373	0.429
		σ	0.127	0.087	0.086	0.093	0.078	0.139	0.049	0.066	0.052	0.073

En resumen, para los tres casos estudiados, en los recursos medidos el método más conservador es SVR, mientras que el menos conservador es ANN. Por otro lado, para los recursos indicados el método más conservador se encuentra entre OCOK y CART, en tanto que el método menos conservador es SVR. El resultado para los cien casos estudiados se discute en la sección 3.3.

Tabla 11. Ranking de la calidad de datos para estimación del mejor, peor y más probable caso según recursos minerales.

		Medidos				Indicados			
		OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
Mejor caso	Cu	4 ^{to}	2	3	1 ^{er}	1 ^{er}	2	3	4 ^{to}
	Mo	2	4 ^{to}	3	1 ^{er}	1 ^{er}	4 ^{to}	2	3
	Au	3	4 ^{to}	1 ^{er}	2	1 ^{er}	3	2	4 ^{to}
Más probable caso	Cu	2	4 ^{to}	3	1 ^{er}	2	3	1 ^{er}	4 ^{to}
	Mo	3	4 ^{to}	2	1 ^{er}	2	4 ^{to}	1 ^{er}	3
	Au	1 ^{er}	4 ^{to}	3	2	2	3	1 ^{er}	4 ^{to}
Peor caso	Cu	3	4 ^{to}	2	1 ^{er}	2	3	1 ^{er}	4 ^{to}
	Mo	1 ^{er}	3	2	4 ^{to}	2	3	1 ^{er}	4 ^{to}
	Au	3	4 ^{to}	2	1 ^{er}	1 ^{er}	3	2	4 ^{to}

3.2.7 ANÁLISIS ECONÓMICO

Como ya ha sido mencionado, se asume que el cobre es el mineral de interés, siendo subproductos el molibdeno y oro. Es por ello que, para simplificar los cálculos, se realiza el análisis económico utilizando la denominada ley equivalente, cuya descripción y ecuaciones se encuentran en el Anexo C. 1. Los datos utilizados para calcular las leyes equivalentes según las ecuaciones (113) y (114) se resumen en la Tabla 12, donde los precios de los metales son obtenidos del día 23 de marzo de 2020 según la Bolsa de Metales de Londres. Conforme a estos datos, la conversión de las leyes de molibdeno y oro al mineral de interés se hacen a partir de la ecuación (80).

$$Ley_{eq} = Ley_{cobre} + 4.3021 \times 10^{-4} \times Ley_{molibdeno} + 0.9699 \times Ley_{oro} \quad (80)$$

Tabla 12. Datos económicos considerados para los casos simulados.

	Cobre	Molibdeno	Oro
Precio	4,617.5 [USD\$/ton]	8.7 [USD\$/lb]	1,554.2 [USD\$/tr oz]
Costo venta [US\$/ton]	1,100	3,200	12,000,000
% pagable de mineral	97	96	96
Recuperación metalúrgica [%]	90	86	82

Suponiendo una densidad de 2.7 [ton/m³], los tonelajes y ley media del cuerpo mineralizado para cada categoría de los recursos minerales en el mejor, peor y más probable caso utilizando la ley

equivalente se resumen en la Tabla 13. En virtud de que los recursos medidos poseen leyes mayores que los recursos indicados, tal como se mostraba anteriormente en la Tabla 10, es que este comportamiento se mantiene luego de desempeñar la ley equivalente. Por otro lado, debido a que la categorización es realizada en base a la varianza de cokriging y replicada en los demás métodos, es que todos comienzan con la misma cantidad de tonelaje en los recursos medidos e indicados, donde para para los tres casos se tiene mayor tonelaje en los recursos medidos, tal como se vio en la Figura 24. Además, es posible apreciar que para los tres casos el método que más se asemeja a la ley media equivalente original total es SVR

Tabla 13. Ley media y tonelaje utilizando ley equivalente para estimación del mejor, peor y más probable caso.

		Ley media [%]					Tonelaje [Mton]				
		Original	OCOK	ANN	CART	SVR	Original	OCOK	ANN	CART	SVR
Mejor caso	Total	1.053	1.043	1.045	1.039	1.053	11.52				
	Medidos	1.178	1.176	1.181	1.165	1.175	7.51				
	Indicados	0.817	0.793	0.790	0.803	0.824	4.01				
Más probable caso	Total	1.134	1.104	1.128	1.098	1.129	11.59				
	Medidos	1.256	1.220	1.256	1.225	1.234	7.85				
	Indicados	0.878	0.859	0.858	0.834	0.909	3.74				
Peor caso	Total	1.088	1.011	1.025	1.018	1.068	11.59				
	Medidos	1.266	1.245	1.264	1.224	1.237	5.86				
	Indicados	0.906	0.771	0.780	0.806	0.896	5.73				

El cálculo de la ley de corte (COG por sus siglas en inglés) se realiza utilizando la ecuación que se encuentran en el Anexo C. 2, donde se considera un costo de mina (CM) de 4.2 [USD\$/ton] y un costo de procesamiento (CP) de 16 [USD\$/ton]. De la ecuación (117) y utilizando los datos del cobre de la Tabla 12, se obtiene una ley de corte de 0.664%, donde en la Tabla 14 se resumen los tonelajes y leyes medias obtenidas para los tres casos analizados después de filtrar la base de datos según esta, haciendo la comparación con las leyes medias obtenidas en la base original utilizando las mismas coordenadas que las de la base estimada después de filtrar. En primera instancia, es posible notar que en el peor caso para el método CART no existen leyes bajo la ley de corte, por lo que este no presenta estéril, sin embargo, para los otros dos casos, el estéril obtenido presenta diferencias mayores en cuanto a su ley media con respecto a la base original. Por otro lado, al tener leyes más altas en los recursos medidos, la cantidad de tonelaje en esta categoría no se ve afectada en gran medida luego de hacer la filtración en las bases de datos estimadas, a diferencia de los recursos indicados donde el

tonelaje disminuye considerablemente. Además, en los tres casos se consigue menor tonelaje en los recursos indicados y mayor tonelaje en el estéril con el método ANN, donde en este último las leyes medias son menores en contraste con los demás. También este método obtiene mayores leyes medias en los recursos medidos, a diferencia de los recursos indicados donde las leyes son más altas con SVR. Por último, es importante señalar que el método SVR es el que más se asemeja a la ley media equivalente total de la base original en los tres casos, tal como ocurre en la base sin filtrar.

Tabla 14. Ley media y tonelaje después de filtrar con ley de corte para estimación del mejor, peor y más probable caso.

COG = 0.664%		Ley media base original [%]				Ley media base estimada [%]				Tonelaje [Mton]			
		OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
Mejor caso	Estéril	0.646	0.647	0.677	0.623	0.622	0.599	0.624	0.619	0.63	0.98	0.75	0.56
	Total	1.076	1.090	1.079	1.075	1.067	1.087	1.068	1.075	10.89	10.54	10.77	10.96
	Medidos	1.178	1.178	1.178	1.178	1.176	1.181	1.165	1.175	7.51	7.51	7.51	7.51
	Indicados	0.849	0.872	0.849	0.849	0.825	0.851	0.844	0.857	3.38	3.02	3.26	3.45
Más probable caso	Estéril	0.663	0.637	0.742	0.665	0.633	0.628	0.657	0.633	0.12	0.43	0.02	0.35
	Total	1.139	1.153	1.134	1.148	1.108	1.147	1.099	1.145	11.47	11.16	11.57	11.24
	Medidos	1.256	1.261	1.256	1.257	1.220	1.262	1.225	1.235	7.85	7.78	7.85	7.84
	Indicados	0.884	0.905	0.878	0.899	0.867	0.882	0.835	0.937	3.63	3.38	3.73	3.40
Peor caso	Estéril	0.707	0.747	–	0.636	0.636	0.635	–	0.649	0.55	1.23	0	0.16
	Total	1.107	1.129	1.088	1.094	1.030	1.071	1.018	1.074	11.04	10.37	11.59	11.44
	Medidos	1.266	1.266	1.266	1.266	1.245	1.264	1.224	1.237	5.86	5.86	5.86	5.86
	Indicados	0.928	0.950	0.906	0.914	0.785	0.820	0.806	0.903	5.18	4.50	5.73	5.57

Para evaluar el beneficio económico, se considera que la extracción de los bloques de mineral del yacimiento se realiza de manera simultánea, sin estimar la extracción adicional que pueda requerir el material estéril. De esta manera, se calcula el ingreso, costo y beneficio para las bases estimadas según la clasificación de recursos para los tres casos ya mencionados, información contenida en la Tabla 15. Se puede observar que para los tres casos se obtienen ingresos y beneficios totales mayores para el método SVR y costos totales menores para ANN. Además, se logran ingresos y beneficios mayores en los recursos indicados con SVR, mientras que para los recursos medidos ambos indicadores son mayores utilizando ANN, siendo este también el método con menores ingresos y costos en los recursos indicados, como era de esperarse según lo comentado de la Tabla 14. De manera general, los ingresos, costos y beneficios son mayores para los recursos medidos en comparación a los recursos indicados.

Tabla 15. Evaluación económica para estimación del mejor, peor y más probable caso.

		Ingreso [MUSD\$]			Costo [MUSD\$]			Beneficio [MUSD\$]		
		Total	Medidos	Indicados	Total	Medidos	Indicados	Total	Medidos	Indicados
Mejor caso	OCOK	353.41	268.59	84.82	220.01	151.73	68.28	133.39	116.86	16.53
	ANN	348.13	269.83	78.30	212.82	151.73	61.08	135.31	118.10	17.22
	CART	349.75	266.18	83.57	217.51	151.73	65.78	132.25	114.45	17.79
	SVR	358.19	268.37	89.81	221.32	151.73	69.59	136.86	116.64	20.22
Más probable caso	OCOK	386.69	291.13	95.56	231.74	158.49	73.25	154.95	132.64	22.31
	ANN	389.18	298.50	90.68	225.41	157.13	68.28	163.76	141.37	22.40
	CART	386.78	292.20	94.57	233.76	158.49	75.27	153.02	133.71	19.31
	SVR	391.33	294.34	96.99	227.05	158.28	68.77	164.28	136.07	28.21
Peor caso	OCOK	345.76	222.09	123.67	223.07	118.46	104.61	122.69	103.62	19.06
	ANN	337.68	225.42	112.26	209.38	118.46	90.92	128.30	106.96	21.34
	CART	358.83	218.35	140.48	234.19	118.46	115.73	124.63	99.89	24.74
	SVR	373.63	220.64	152.99	231.03	118.46	112.57	142.60	102.18	40.42

La Tabla 16 contiene la diferencia entre los beneficios obtenidos en la base original y los beneficios obtenidos en la base estimada según las categorías de los recursos en los tres casos (ver Anexo D. 2 para las diferencias de beneficio económico en las cien simulaciones). Si bien no todas las diferencias menores referidas en la categoría de los recursos son obtenidas por el método SVR, este sí la alcanza en la diferencia total en los tres casos. Por otro lado, un mayor valor numérico en la diferencia total es obtenido por el método CART, a excepción del peor caso, donde este es superior para OCOK. En cuando a los recursos, para el mejor y peor caso, se obtienen mayores diferencias en los indicados y en el caso más probable ocurre lo contrario.

Tabla 16. Diferencia entre los beneficios de la base original y la base estimada para el mejor, peor y más probable caso.

	Mejor caso [MUSD\$]			Más probable caso [MUSD\$]			Peor caso [MUSD\$]		
	Total	Medidos	Indicados	Total	Medidos	Indicados	Total	Medidos	Indicados
OCOK	2.97	0.53	2.44	10.56	8.60	1.96	26.06	3.64	22.43
ANN	1.22	-0.71	1.92	2.11	-0.24	2.34	18.05	0.30	17.75
CART	3.48	2.94	0.54	12.45	7.53	4.92	24.83	7.37	17.46
SVR	-0.15	0.75	-0.90	1.22	5.17	-3.95	6.99	5.08	1.91

3.2.7.1 Selección de casos

Se selecciona el mejor, peor y más probable caso para ser comparados ya que estos son los escenarios más interesantes. El mejor (peor) caso elegido se selecciona para la simulación que obtengan menores (mayores) valores en la diferencia del beneficio de la base original y el beneficio de la base estimada para los métodos utilizados, cuyo procedimiento se especifica a continuación:

1. Se calcula la diferencia del beneficio económico entre la base original y las bases estimadas con los cuatro métodos utilizados para las cien simulaciones.
2. Las diferencias obtenidas son normalizadas para una mejor comparación entre datos.
3. Para identificar cuál de las simulaciones obtuvo mejores (o peores) resultados, se realiza un promedio de los resultados obtenidos en el punto 2 entre los cuatro métodos, obteniendo valores entre 0 y 1, donde 0 es el mejor caso y 1 es el peor.

De esta manera, se seleccionan los casos donde la cercanía o lejanía en cuanto al beneficio económico real sea simultáneo para los cuatro métodos de estimación utilizados. Finalmente, el mejor, peor y más probable caso son las simulaciones 62, 31 y 7, respectivamente.

3.3 ANÁLISIS DE RESULTADOS

Dado lo evidenciado en los tres casos estudiados en cuanto a medidas de error de estimación e incertidumbre económica, se tiene una presunción de obtener mejores resultados utilizando el método *Support Vector Regression*, contrario a lo que ocurre con los Árboles de Regresión, que mostraron un rendimiento inferior. La Tabla 17 comprende la cantidad de veces en las que un método de estimación obtiene mejor y peor rendimiento en cuanto al *skill value* en comparación con el resto, en la que se distingue que el método SVR es mejor que los demás el 86% de las veces, siendo este porcentaje menor cuando se trata de ANN y OCOK, que son mejores el 10.3 y 3.7% de las veces, respectivamente y, por último, en cuando al método CART, este nunca obtiene mejores resultados. Por otro lado, CART consigue resultados inferiores que los otros métodos en promedio un 79.7% de las ocasiones, siendo lo contrario para SVR que en general nunca es peor que los demás.

Tabla 17. Porcentaje de veces en las que el método de estimación utilizado obtiene mejor y peor *skill value*.

	Veces en las que el método obtiene mejor				Veces en las que el método obtiene peor			
	<i>skill value</i>				<i>skill value</i>			
	OCOK [%]	ANN [%]	CART [%]	SVR [%]	OCOK [%]	ANN [%]	CART [%]	SVR [%]
Cobre	5	8	0	87	12	8	79	1
Molibdeno	4	10	0	86	11	13	76	0
Oro	2	13	0	85	9	7	84	0

Teniendo en cuenta la hipótesis de este trabajo de investigación, donde se quiere comprobar que los métodos de *Machine Learning* obtienen mejores resultados que los métodos tradicionales de geostatística ante una baja correlación espacial y alta correlación entre variables, es que se hace la comparación entre los métodos de IC y el método Cokriging Ordinario. En la Tabla 18 se especifican la cantidad de veces en que los métodos de estimación utilizando ML son mejores en comparación al OCOK en base al *skill value*, donde se evidencia que el SVR obtiene mayores porcentajes de superioridad en contraste con el Cokriging, ganándole en promedio el 95.3% de las veces. El método siguiente es ANN ganando aproximadamente el 65.3% de las veces y, por último, se encuentra CART con porcentajes más bajos, ganando alrededor del 13% de las ocasiones.

Tabla 18. Porcentaje de veces en las que ML obtiene mejor *skill value* que Cokriging Ordinario.

	ANN – OCOK [%]	CART – OCOK [%]	SVR – OCOK [%]
Cobre	62	13	93
Molibdeno	69	14	96
Oro	65	12	97

Para dimensionar que tan buenos o malos son los métodos utilizados, se construye la Tabla 19 que indica el promedio de los porcentajes de mejora, para los casos en que los métodos de ML son superiores, y empeoramiento, en los que ML es inferior en comparación con el método tradicional en las cien simulaciones, ambas en función de los *skill values* obtenidos. Es distinguible el porcentaje de mejora que presenta SVR en comparación con ANN y CART, que es alrededor de un 6.81%, donde, por el contrario, el 4.7% de las veces en las que OCOK obtiene un *skill value* mayor a este, su rendimiento disminuye en un 1.76%. Por otro lado, CART, que es el método que supera a OCOK en un menor número de veces, alcanza una mejora del 1.71%, sin embargo, cuando Cokriging obtiene mejores resultados, su rendimiento disminuye en un 7.1%.

Tabla 19. Promedio de porcentaje de mejoramiento y empeoramiento de ML versus Cokriging Ordinario.

	Casos en los que es mejor ML vs OCOK			Casos en los que es mejor OCOK vs ML		
	ANN [%]	CART [%]	SVR [%]	ANN [%]	CART [%]	SVR [%]
Cobre	↑ 4.43	↑ 1.94	↑ 6.93	↓ 3.55	↓ 7.44	↓ 1.68
Molibdeno	↑ 4.79	↑ 1.26	↑ 7.02	↓ 4.20	↓ 6.34	↓ 2.40
Oro	↑ 4.12	↑ 1.93	↑ 6.49	↓ 3.90	↓ 7.42	↓ 1.22

En cuanto a la calidad de los datos según la categorización de recursos, se define como α a la diferencia entre la ley media de la base original y la ley media de la base estimada y como β a la desviación estándar de la base estimada, ambas calculadas para una categorización determinada según número de simulación. La Tabla 20 contiene el promedio de α ($\bar{\alpha}$) y de β ($\bar{\beta}$) entre las simulaciones según la categoría de recurso mineral, donde se puede apreciar que, en los recursos medidos, el número de veces en las que la ley media de la base original es menor que la ley media de la base estimada ($\alpha < 0$) son mayores para Cokriging Ordinario y Redes Neuronales, es decir, estiman leyes más altas, mientras que el número de veces en las que la ley media de la base original es mayor que la ley media de la base estimada ($\alpha > 0$) es mayor para *Support Vector Regression*, estimando leyes más bajas. Por otro lado, en los recursos indicados ocurre todo lo contrario, ya que son más las veces en las que $\alpha < 0$ para SVR, mientras que las ocasiones cuando $\alpha > 0$ son mayores para OCOK y ANN. Esto permite tener una idea de qué tan conservadores son los métodos, pues se puede deducir que SVR es más conservador en los recursos medidos y menos conservador en los indicados, en tanto que para OCOK y ANN es lo opuesto, siendo más arriesgados en los medidos y más conservadores en los indicados. Además, tanto para los recursos medidos e indicados, el método SVR obtiene casi siempre un valor más cercano a la ley media real ($\bar{\alpha}$) en las tres variables continuas, siendo lo contrario para Árboles de Regresión, que el valor siempre se aleja más. También se observa que cuando $\alpha > 0$, el promedio de la desviación estándar de ANN es mayor a los demás métodos, tanto para los recursos medidos como indicados, siendo las leyes más dispersas al ser la ley media menor a la de la base original, mientras que cuando $\alpha < 0$, el promedio de la desviación estándar es casi siempre menor para SVR, es decir, la variación de las leyes es más pequeña cuando la ley media es mayor a la ley media de la base real.

Tabla 20. Promedio de la calidad de los datos estimados en las cien simulaciones según categoría de recursos.

		Cobre [%]				Molibdeno [ppm]				Oro [ppm]				
		OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	
Medidos	$\alpha < 0$	Veces	53	47	40	20	42	56	31	16	45	47	34	22
	$ \bar{\alpha} $	1.225	1.117	1.194	0.467	1.140	0.962	1.175	0.467	1.012	1.107	1.102	0.342	
	$\bar{\beta}$	8.450	8.888	8.797	8.615	8.761	8.658	8.794	8.626	8.640	8.587	8.625	8.562	
	$\alpha > 0$	Veces	47	53	60	80	58	44	69	84	55	53	66	78
	$\bar{\alpha}$	1.108	1.003	1.758	0.926	1.130	0.917	1.603	0.867	1.083	0.900	1.528	0.828	
	$\bar{\beta}$	7.371	8.456	7.653	7.821	8.208	8.774	8.005	8.008	8.140	9.119	8.168	8.117	
Indicados	$\alpha < 0$	Veces	30	37	58	66	48	45	58	77	50	42	56	73
	$ \bar{\alpha} $	1.227	1.088	2.130	1.102	1.547	1.208	2.165	1.219	1.514	1.242	2.331	1.205	
	$\bar{\beta}$	6.545	6.570	6.868	6.190	6.680	6.647	6.954	6.382	6.853	7.237	6.773	6.500	
	$\alpha > 0$	Veces	70	63	42	34	52	55	42	23	50	58	44	27
	$\bar{\alpha}$	1.725	1.347	1.881	0.560	1.564	1.251	1.799	0.537	1.683	1.284	1.922	0.427	
	$\bar{\beta}$	5.997	6.995	6.226	6.868	6.795	7.294	6.455	6.812	6.653	6.967	6.365	6.854	

Con el fin de determinar qué tan conservadores o arriesgados son los métodos, se crea un ranking con la información contenida en la Tabla 20 y siguiendo la metodología presentada en la sección 3.2.6, determinando el número de veces en las que un método es más o menos conservador según la clasificación de recursos, lo que es comprendido en la Tabla 21. En esta se aprecia que en el caso de los recursos medidos el método que es más veces conservador que el resto es SVR, mientras que el más arriesgado es Redes Neuronales. En tanto, para los recursos indicados ocurre que el método más conservador es Cokriging Ordinario y el menos conservador es CART.

Tabla 21. Porcentaje de veces en la que un método es más o menos conservador según categoría de recursos.

		Veces en las que el método es más conservador			Veces en las que el método es menos conservador		
		Cobre [%]	Molibdeno [%]	Oro [%]	Cobre [%]	Molibdeno [%]	Oro [%]
Medidos	OCOK	6	6	15	24	29	19
	ANN	15	8	12	48	52	53
	CART	32	35	33	16	11	14
	SVR	47	51	40	12	8	14
Indicados	OCOK	47	31	23	1	15	13
	ANN	24	26	29	17	22	21
	CART	18	28	27	48	35	33
	SVR	11	15	21	34	28	33

Ahora bien, para tener una idea más general de la calidad de los datos es que se entrega un ranking comparando los parámetros antes mencionados, α y β , para las bases estimadas totales sin diversificar por recursos. Según la Tabla 22, que contiene el porcentaje de las veces en las que el método es más o menos conservador, el método más conservador se encuentra entre SVR y CART, en tanto que el más arriesgado es evidentemente ANN. Es interesante notar que para el OCOK existe una tendencia a aumentar el porcentaje de ser conservador a medida que la variable continua resulta ser más errática, ocurriendo lo contrario para SVR, que se vuelve menos conservador a medida que esto pasa.

Tabla 22. Porcentaje de veces en las que un método es más o menos conservador en general.

	Veces en las que el método es más conservador			Veces en las que el método es menos conservador		
	Cobre [%]	Molibdeno [%]	Oro [%]	Cobre [%]	Molibdeno [%]	Oro [%]
OCOK	9	10	17	25	26	16
ANN	20	14	20	46	51	54
CART	35	35	35	20	13	17
SVR	36	41	28	9	10	13

Para tener una mejor visión de los resultados de la evaluación económica es que se determina el número de veces en las que un método se asemeja más al beneficio económico de la base original y el número de veces en las que más se diferencia, además de mostrar el porcentaje de error del beneficio estimado en contraste del obtenido realmente. La Tabla 23 muestra que el 59% de las ocasiones el método SVR obtiene valores más cercanos al valor real del beneficio económico, donde solo el 2% de las veces el error cometido es mayor. Por otro lado, CART es el método que produce un beneficio económico más alejado de la realidad con un 58% de las veces, mientras que un 5% de los casos el valor obtenido es más similar al original. En cuanto a los errores producidos, SVR los produce en un menor porcentaje, con un margen del $\pm 2.38\%$, mientras que el margen mayor es de $\pm 7.09\%$ provocado por Árboles de Regresión.

Por último, se hace la comparación de las veces en que son mejores los métodos de ML en comparación con Cokriging Ordinario, es decir, que el error cometido en una misma simulación es menor para los primeros mencionados, información comprendida en la Tabla 24. Los métodos SVR

y ANN son superiores al OCOK la mayoría de las veces, con un 84 y 62%, respectivamente, mientras que, por el contrario, el método CART es inferior la mayor parte de las ocasiones, con un 72%.

Tabla 23. Porcentaje de veces en las que un método se asemeja o aleja del beneficio económico real.

	Porcentaje de veces [%]		Porcentaje de error [%]
	Mejor resultado	Peor resultado	
OCOK	8	19	5.52
ANN	28	21	4.37
CART	5	58	7.09
SVR	59	2	2.38

Tabla 24. Porcentaje de veces en que ML es mejor que OCOK en cuanto al error producido en el beneficio económico.

Veces en las que ML es mejor que OCOK	
ANN – OCOK [%]	62
CART – OCOK [%]	28
SVR – OCOK [%]	84



CAPÍTULO IV: CASO REAL

4.1 METODOLOGÍA DE TRABAJO

El procedimiento de trabajo empleado para la base con datos reales es el siguiente:

- **Estudio exploratorio de la base de datos:** se realiza una descripción y análisis completo de la base de datos, reportando estadísticas principales de las variables contenidas en esta, donde además se realiza una limpieza de datos en caso de ser requerido.
- **Creación del modelo de bloques:** se generan nuevas coordenadas donde hacer la estimación de las variables continuas contenidas en la base de datos.
- **Particionamiento de base:** la base es dividida en dos conjuntos, uno que contiene los datos para entrenar los algoritmos y otro que contiene los datos para hacer la validación de la estimación, con el objetivo de entrenarlas con el menor error de estimación.
- **Métodos de estimación:** se estiman todas las variables continuas presentes en la base de datos en las coordenadas de la base de validación para buscar una combinación óptima de parámetros en cada uno de los cuatro algoritmos utilizados, donde luego se realiza la estimación en el nuevo modelo de bloques.
- **Medidas de error de estimación:** se utilizan las medidas de error ME, MAE, RMSE y el coeficiente de determinación R^2 resumidos en el *skill value* para realizar un estudio comparativo entre los métodos utilizados para cada una de las variables continuas presentes en el caso de estudio.
- **Medidas estadísticas principales del modelo de bloques:** se reportan algunas medidas estadísticas de las leyes de los minerales presentes estimados en el modelo de bloques.
- **Categorización de recursos minerales:** se hace la categorización de recursos minerales utilizando la varianza arrojada por el Cokriging Ordinario para el mineral de interés, reportando estadísticas principales tanto para los datos estimados en el nuevo modelo de bloques como en la base de validación, donde en esta última se determina la calidad de los datos.
- **Análisis económico:** se realiza una evaluación económica para las cuatro bases estimadas en el modelo de bloques.

4.2 ESTUDIO EXPLORATORIO DE DATOS

La base de datos real cuenta con 7997 observaciones de un yacimiento bimetálico que contiene 6 variables, las que corresponden a las coordenadas *Easting*, *Northing* y *Elevation* expresadas en UTM, leyes de cobre y oro medidas en porcentaje [%] y partes por millón [ppm] respectivamente, y, por último, una variable que describe cuatro dominios geometalúrgicos presentes en el cuerpo mineralizado. La distribución de las muestras en el espacio se encuentra en una configuración irregular, donde la zona cubierta es de aproximadamente 1.8 km al este, 1.3 km al norte y 113 m en dirección vertical, según lo contenido en la Tabla 25. El yacimiento presenta leyes medias globales de 0.352% y 0.320 ppm para el cobre y oro, respectivamente, mientras que las leyes medias según el dominio se resumen en la Tabla 26 junto con otras medidas estadísticas. Según la tabla anteriormente mencionada, las leyes medias, tanto para el cobre como para el oro, son mayores en el dominio 3 y estas, a su vez, son más dispersas que en los otros dominios, en especial en el caso del oro, mientras que en el dominio 2 las leyes medias son menores. Por otro lado, el dominio 3 es el que contiene menos observaciones con un 11.3% de la base total, en tanto que el dominio 4 es que el posee mayor cantidad de datos, con un 32.2% de la base total.

Tabla 25. Coordenadas máximas y mínimas de la base de datos real.

	<i>Easting</i>	<i>Northing</i>	<i>Elevation</i>
Mínimo	674,036.2	424,650.9	255
Máximo	675,832.4	425,907.7	367.5

Tabla 26. Medidas estadísticas principales de las leyes de cobre y oro para cada dominio de la base real.

	Cobre [%]				Oro [ppm]			
	Dominio 1	Dominio 2	Dominio 3	Dominio 4	Dominio 1	Dominio 2	Dominio 3	Dominio 4
Media	0.276	0.266	0.708	0.369	0.215	0.212	0.731	0.364
Desviación estándar	0.181	0.181	0.264	0.155	0.162	0.183	0.447	0.204
Mediana	0.240	0.224	0.730	0.360	0.170	0.156	0.718	0.330
Mínimo	0.001	0.004	0.038	0.003	0	0	0.052	0.009
Máximo	0.880	0.998	1.720	0.955	1.170	1.130	3.700	1.110
N° de datos	2,166	2,349	903	2,579	2,166	2,349	903	2,579

En la Figura 25 se encuentra un gráfico de caja de la distribución de las leyes de cobre y oro según la base total y sus dominios, donde efectivamente se comprueba que el dominio 3 es el que contiene mayores leyes en ambos casos, seguido por el dominio 4. Además, se visualizan los probables datos atípicos para ambos minerales en cada dominio, en especial en el caso del oro para el dominio 3 donde se observan tres datos muy distantes del resto.

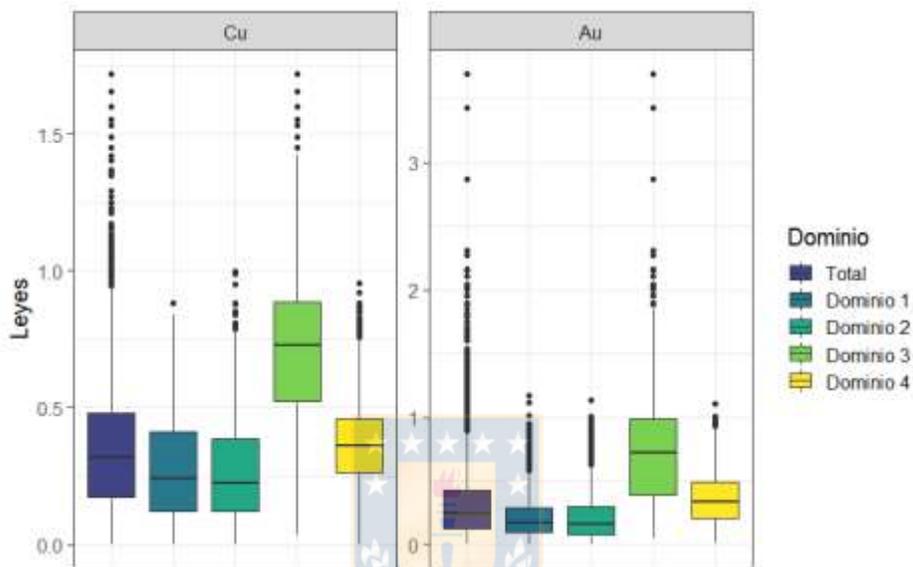


Figura 25. Boxplot de las leyes de cobre y oro para la base real en general y según dominios.

Es relevante mencionar que la correlación entre las leyes de cobre y oro es en general de 0.78. Ahora bien, la correlación por dominio es de 0.75, 0.81, 0.61 y 0.63 para el dominio uno, dos, tres y cuatro, respectivamente. Finalmente, la visualización de ambas leyes distribuidas en el espacio y los dominios geometalúrgicos presentes en el yacimiento se visualizan en la Figura 26, donde se observa que el dominio 3 se encuentra en gran medida a mayor profundidad a diferencia de los otros tres.

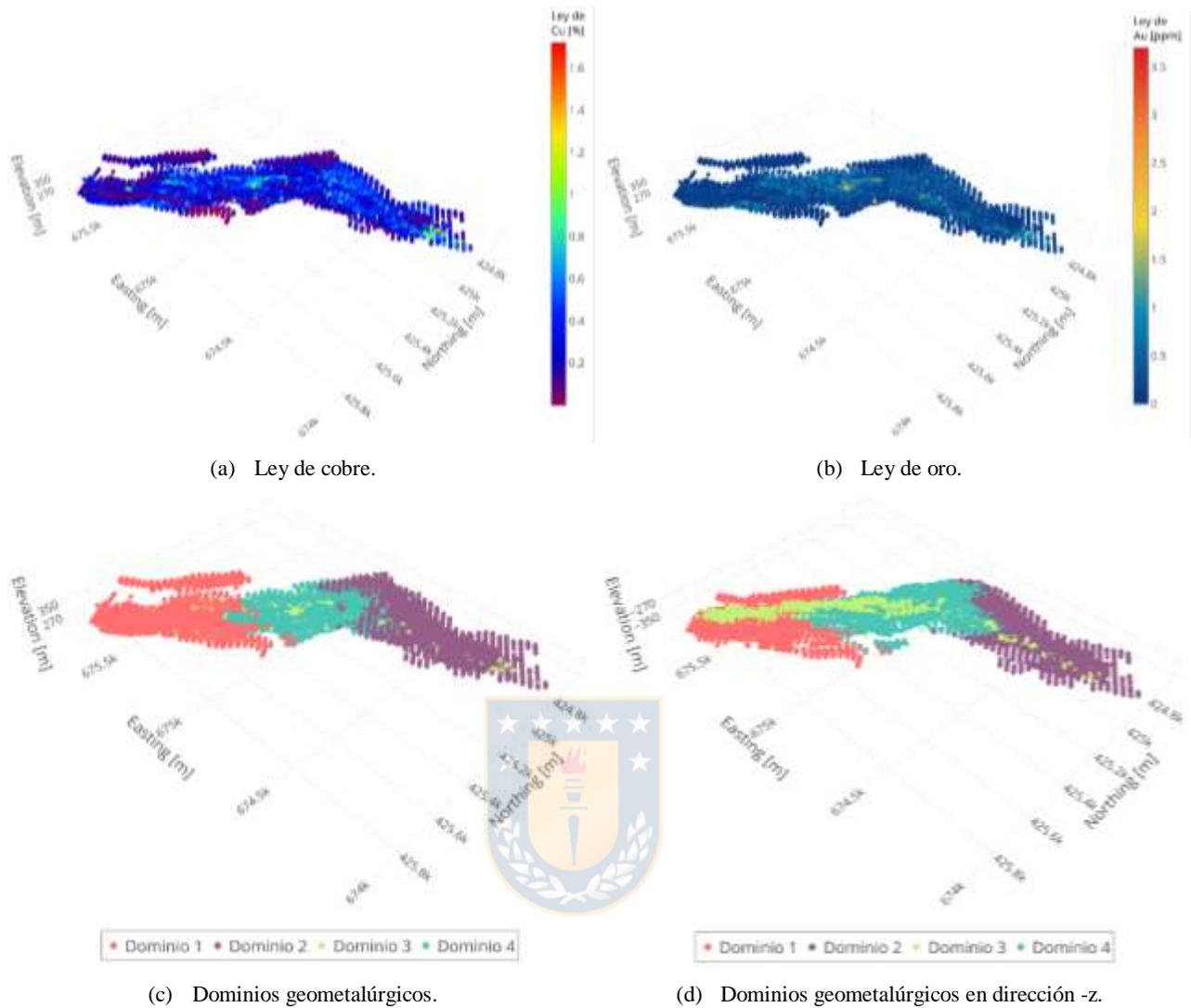


Figura 26. Visualización de leyes de cobre y oro, además de los dominios presentes en la base real.

4.2.1 DATOS DUPLICADOS Y OUTLIERS

Un dato duplicado es considerado aquel que en una misma coordenada tiene más de una medición para una misma variable, de manera que no se sabe si hubo error de tipeo o la información contenida es errónea. Es por esta razón que una coordenada con información repetida es eliminada, ya que no se tiene certeza de la veracidad de esta. Analizando la base de datos entregada, se encuentran siete coordenadas que cuentan con doble observación para las leyes de cobre y oro, resultando en la eliminación de 14 datos.

Por otro lado, un dato aberrante, o también llamado *outlier*, es una observación atípica que es numéricamente inconsistente con el resto. Como se cuenta con cuatro dominios geometalúrgicos, se hace un análisis en cada uno de ellos para determinar los datos atípicos existentes utilizando la ecuación (81) para las leyes de cobre y la ecuación (82) para el oro.

$$Ley_{Cu} < \overline{Ley_{Cu}} - 3.5 \times \sigma_{Ley_{Cu}} \mid Ley_{Cu} > \overline{Ley_{Cu}} + 3.5 \times \sigma_{Ley_{Cu}} \quad (81)$$

$$Ley_{Au} < \overline{Ley_{Au}} - 4.0 \times \sigma_{Ley_{Au}} \mid Ley_{Au} > \overline{Ley_{Au}} + 4.0 \times \sigma_{Ley_{Au}} \quad (82)$$

Es importante señalar que si bien a veces los datos atípicos no se deben a algún error de medida y no deben ser removidos o sustituidos sobre todo en el caso del oro, se decide eliminarlos ya que después de varios intentos en la estimación de las variables continuas del yacimiento y aprendiendo de las pruebas y sus errores, es que se llega a la conclusión que las leyes muy altas en la base de datos estudiada, en especial las del dominio 3, conllevan a un alto error en la estimación de estas.

Son 8 los datos calificados como *outliers* para el caso de las leyes de cobre y 22 para las leyes de oro, siendo un total de 30 observaciones atípicas, lo que suma 44 datos eliminados. La base de datos luego de ser limpiada consta de 7953 observaciones, donde las leyes medias generales de cobre y oro disminuyen levemente a 0.351% y 0.317 ppm, respectivamente, mientras que las medidas estadísticas principales según dominio se encuentran en la Tabla 27. Además, en la Figura 27 se visualiza el gráfico de caja de las leyes minerales estudiadas después de la limpieza, donde es notable la diferencia de los *outliers* apreciables en la Figura 25.

Tabla 27. Medidas estadísticas principales de la base real según dominio después de limpiarla.

	Cobre [%]				Oro [ppm]			
	Dominio 1	Dominio 2	Dominio 3	Dominio 4	Dominio 1	Dominio 2	Dominio 3	Dominio 4
Media	0.276	0.264	0.706	0.369	0.212	0.208	0.721	0.364
Desviación estándar	0.181	0.179	0.260	0.154	0.155	0.175	0.420	0.204
Mediana	0.240	0.224	0.730	0.360	0.169	0.155	0.716	0.330
Mínimo	0.001	0.004	0.038	0.003	0	0	0.052	0.009
Máximo	0.880	0.880	1.600	0.883	0.860	0.942	2.310	1.110
N° de datos	2.158	2.331	898	2.566	2.158	2.331	898	2.566

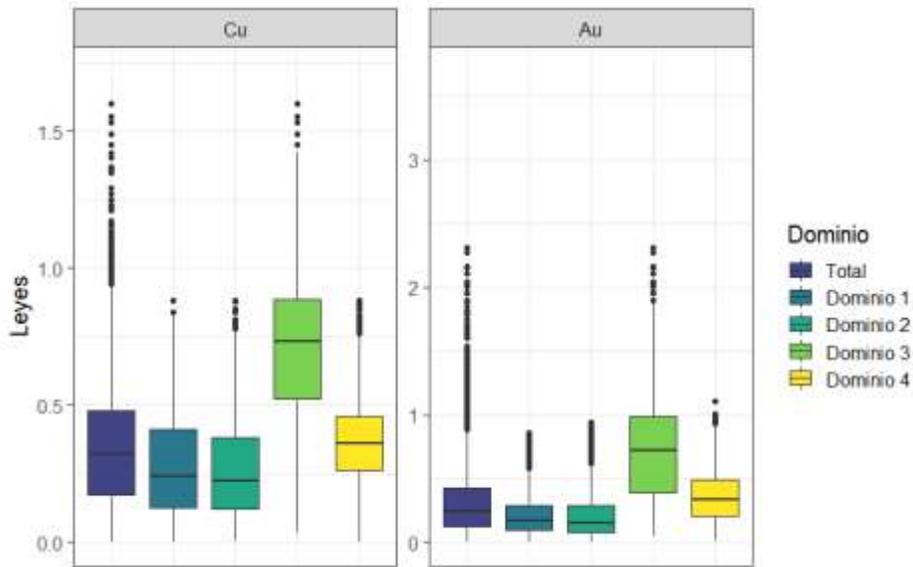


Figura 27. Boxplot de las leyes de cobre y oro para la base real en general y según dominios después de limpiarla.

Gracias a la limpieza realizada, la correlación entre las leyes de cobre y oro de manera global aumentaron a 0.80, mientras que por dominio aumentaron a 0.78, 0.83, 0.65 y 0.64 para el dominio uno, dos, tres y cuatro, respectivamente.



4.3 APLICACIÓN

4.3.1 CREACIÓN DEL MODELO DE BLOQUES

El espacio de trabajo en donde se van a realizar las estimaciones consta de una extensión de las coordenadas presentes en la base de datos real, manteniendo de esta manera la forma del yacimiento. En este caso, los valores mínimos y máximos creados en el modelo de bloques se encuentran en la Tabla 28, cubriendo una zona aproximadamente de 1.96 km al este, 1.38 km al norte y 120 m en dirección vertical. Se tienen un total de 51,784 bloques, donde cada uno tiene un tamaño de 20 x 20 x 10 [m³], visualizado en la Figura 28.

Tabla 28. Coordenadas máximas y mínimas del modelo de bloques.

	<i>Easting</i>	<i>Northing</i>	<i>Elevation</i>
Mínimo	673,946.4	424,588.1	249.375
Máximo	675,906.4	425,968.1	369.375

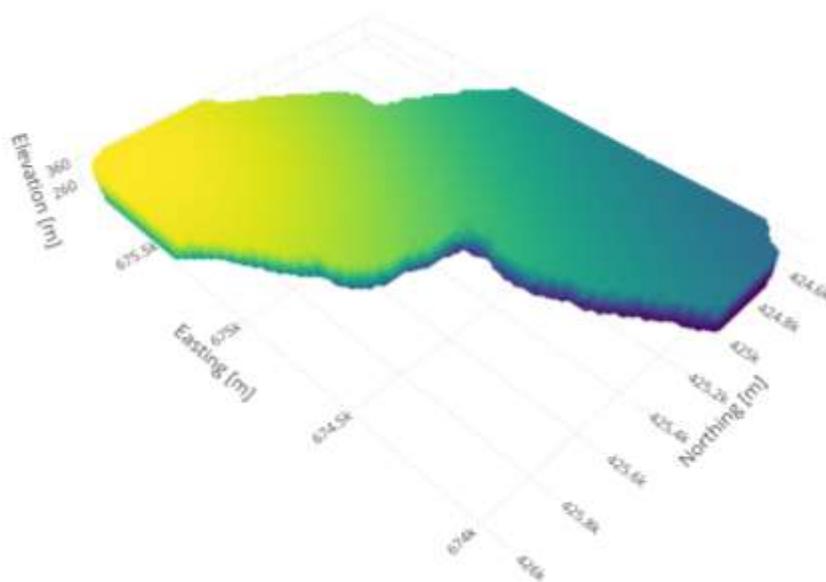


Figura 28. Modelo de bloques para la base real.

4.3.1.1 Extensión de dominios geometalúrgicos

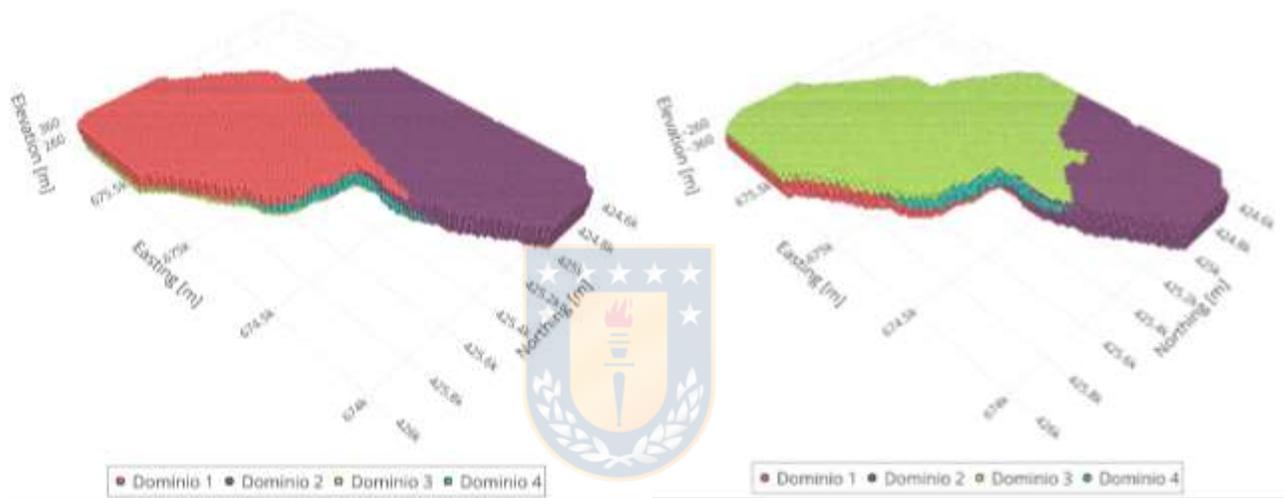
Es necesario tener presente el tiempo que un algoritmo demora en procesar los datos, donde, considerando que cada uno de ellos se comportan de forma diferente, existen algunos más rápidos y otros bastante más lentos. Dada la cantidad de datos contenida en la base real y en el modelo de bloques creado es que, se decide realizar las estimaciones con cada método utilizado según el dominio geometalúrgico, ya que de esa manera el tiempo en el procesamiento de los datos será menor, siendo esencial extenderlos en el modelo de bloques creado.

Para realizar la extensión de los dominios se utiliza el método SVM para clasificación con un kernel Gaussiano, utilizando un 60% de la base de datos para entrenar el modelo, mientras que con el 40% restante se hace la validación. Se optimiza el parámetro de regularización C , γ del kernel utilizado según la ecuación (73) y la tolerancia de manera simultánea en función de la precisión, creando una grilla donde los valores de los parámetros a ser optimizados son $C = \{2^{-5}, 2^{-4.5}, \dots, 2^7\}$, $\gamma = \{2^{-7}, 2^{-6.5}, \dots, 2^2\}$ y $\text{tolerancia} = \{0.05, 0.1, 0.12, 0.15, 0.18\}$, obteniendo un total de 2375 combinaciones. Finalmente, los valores utilizados para extender los dominios en el modelo de bloques son $C = 2^6$, $\gamma = 2^{-4.5}$ y una tolerancia de 0.15, obteniendo una precisión o porcentaje de aciertos del 99.4% obtenida en la base real según la Tabla 29, la cual contiene la cantidad de datos acertados y

precisión para cada dominio. En la Figura 29 en encuentra la visualización del modelo de bloques con los dominios obtenidos en la predicción.

Tabla 29. Cantidad de aciertos en la predicción de dominios para la base real.

	Dominio 1	Dominio 2	Dominio 3	Dominio 4	Total	Precisión [%]
Dominio 1	2145	0	5	8	2158	99.4
Dominio 2	0	2323	1	7	2331	99.7
Dominio 3	0	0	888	10	898	98.9
Dominio 4	2	14	1	2549	2566	99.3



(a) Dominios en dirección z+.

(b) Dominios en dirección z-.

Figura 29. Extensión de dominios geometalúrgicos en modelo de bloques.

4.3.2 PARTICIONAMIENTO DE BASE

Al tener presente la premisa de trabajar con poca información es que la división de la base de datos para el caso real se hace considerando un 60% de cada dominio para realizar el entrenamiento de los algoritmos, lo que, en contraste con el modelo de bloques, corresponde a un 9.2% de datos, mientras que el 40% restante es utilizado para hacer la validación y optimización de los parámetros utilizados en cada método. Además, como ya fue mencionado, es importante mantener las medidas estadísticas principales de la base original en las nuevas dos bases creadas para obtener un buen resultado en la estimación, siendo el método ANOVA el utilizado para garantizar esta condición. En la Figura 30 se visualiza un gráfico de caja y de densidad para comparar la distribución entre la base

de datos real y las bases generadas en la partición, donde se comprueba que las estadísticas principales son similares entre ellas. De esta manera, la base de entrenamiento contiene 4773 datos, en tanto que la base de validación presenta 3180 observaciones.

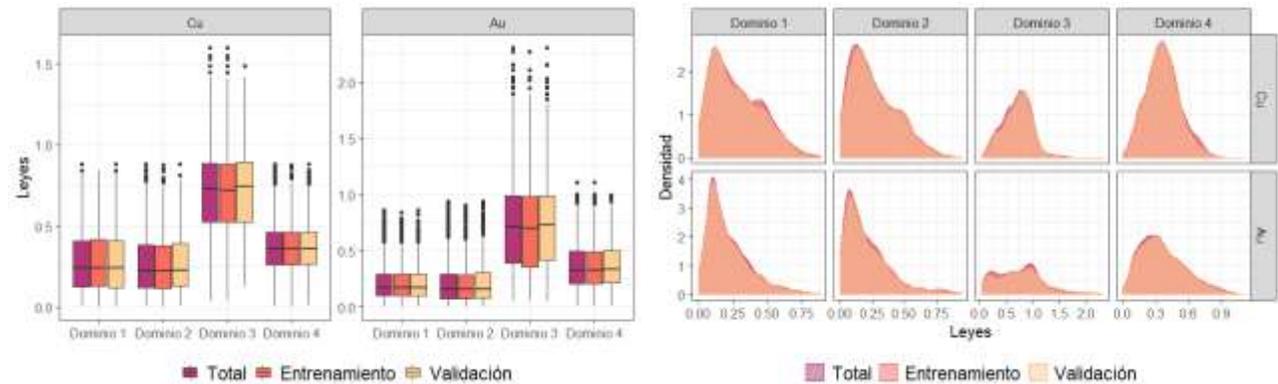


Figura 30. Distribución de la base de datos real, de entrenamiento y validación.

Finalmente, la correlación entre las variables continuas, ley de cobre y oro, de manera global es de 0.79, mientras que por dominio es de 0.77, 0.84, 0.61 y 0.63 para el dominio 1, 2, 3 y 4, respectivamente, siendo en general similares a las correlaciones en la base total.

4.3.3 MÉTODOS DE ESTIMACIÓN

A continuación, se presenta la metodología adquirida en la estimación para cada uno de los métodos utilizados en el trabajo de investigación, donde primero se hace la estimación en la base de validación para cada dominio y, una vez optimizados los parámetros de cada método, se estima en el modelo de bloques creado.

4.3.3.1 Cokriging Ordinario

Tal como ya fue mencionado, el primer paso para realizar una estimación con los métodos de kriging y sus variantes es revisar el mapa variográfico con el objetivo de determinar la existencia de anisotropía, para luego estudiar los variogramas direccionales en caso de ser necesario. Como la estimación es realizada según el dominio geometalúrgico, es que se estudia el mapa variográfico en cada uno de ellos como se observa en la Figura 31, donde es evidente la existencia de anisotropía para

el dominio 1 y 2 en una dirección de mayor continuidad con orientación aproximada de 45° y 70° , respectivamente, ambas medidas en sentido horario desde el norte, mientras que para el dominio 3 y 4 se observa una tendencia más leve que para los dominios anteriores, con una dirección principal de 45° . Esto se confirma al estudiar los variogramas direccionales en cada dominio, pues parece ser que la estructura espacial es relativamente más fuerte en las direcciones principales mencionadas que en las direcciones ortogonales, presentando una variabilidad menor, tal como se visualiza en la Figura 32. Además, en dicha imagen se puede observar que los valores del variograma para el dominio 3 son mayores que para los otros dominios, por lo que es esperable que se obtengan errores superiores. Por otro lado, también es distinguible que para el dominio 3 y 4 las mesetas de los variogramas son bastante cercanas a las pepas, demostrando que la correlación espacial es menor que en los otros dos dominios.

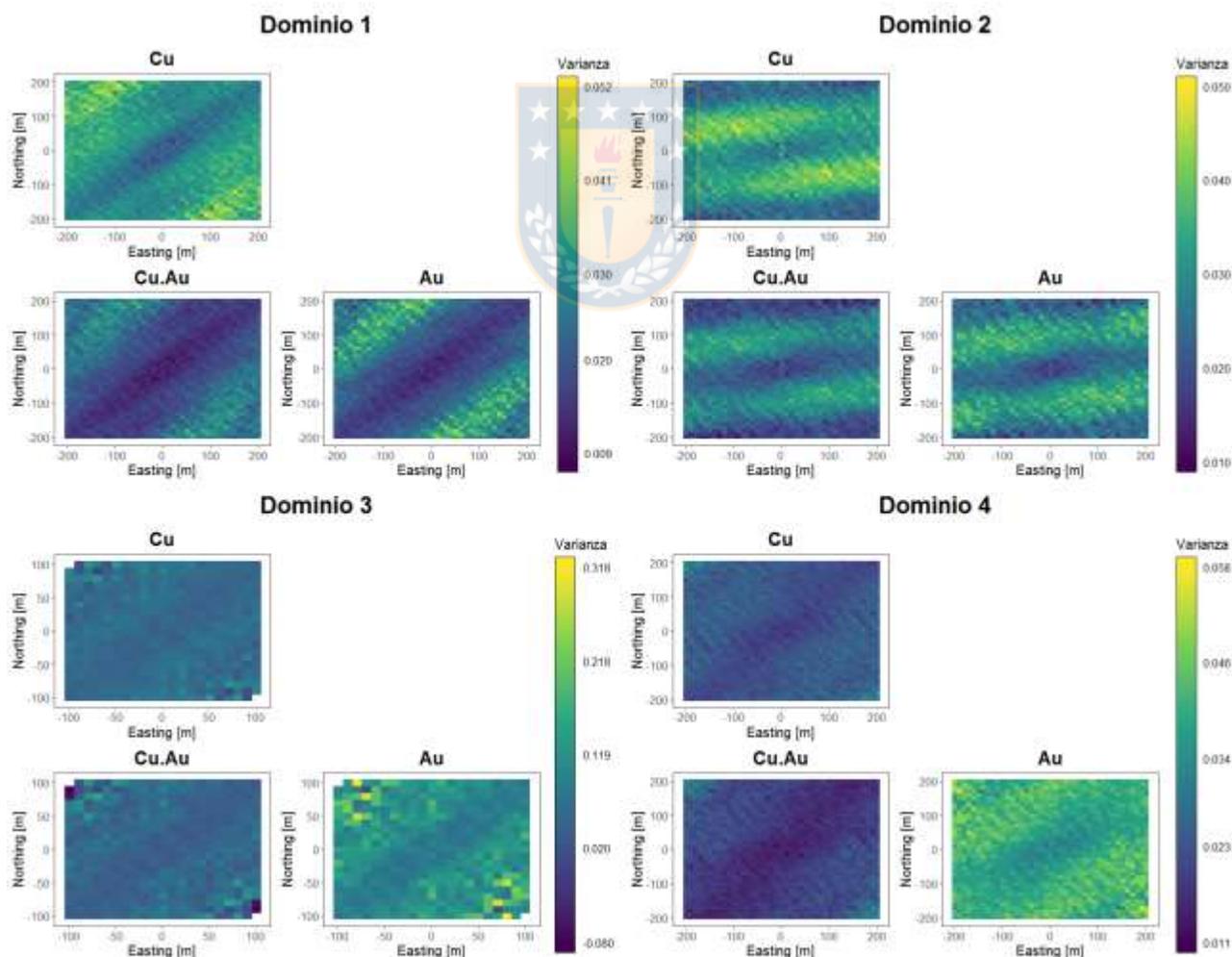


Figura 31. Mapas variográficos de la base real según dominio.

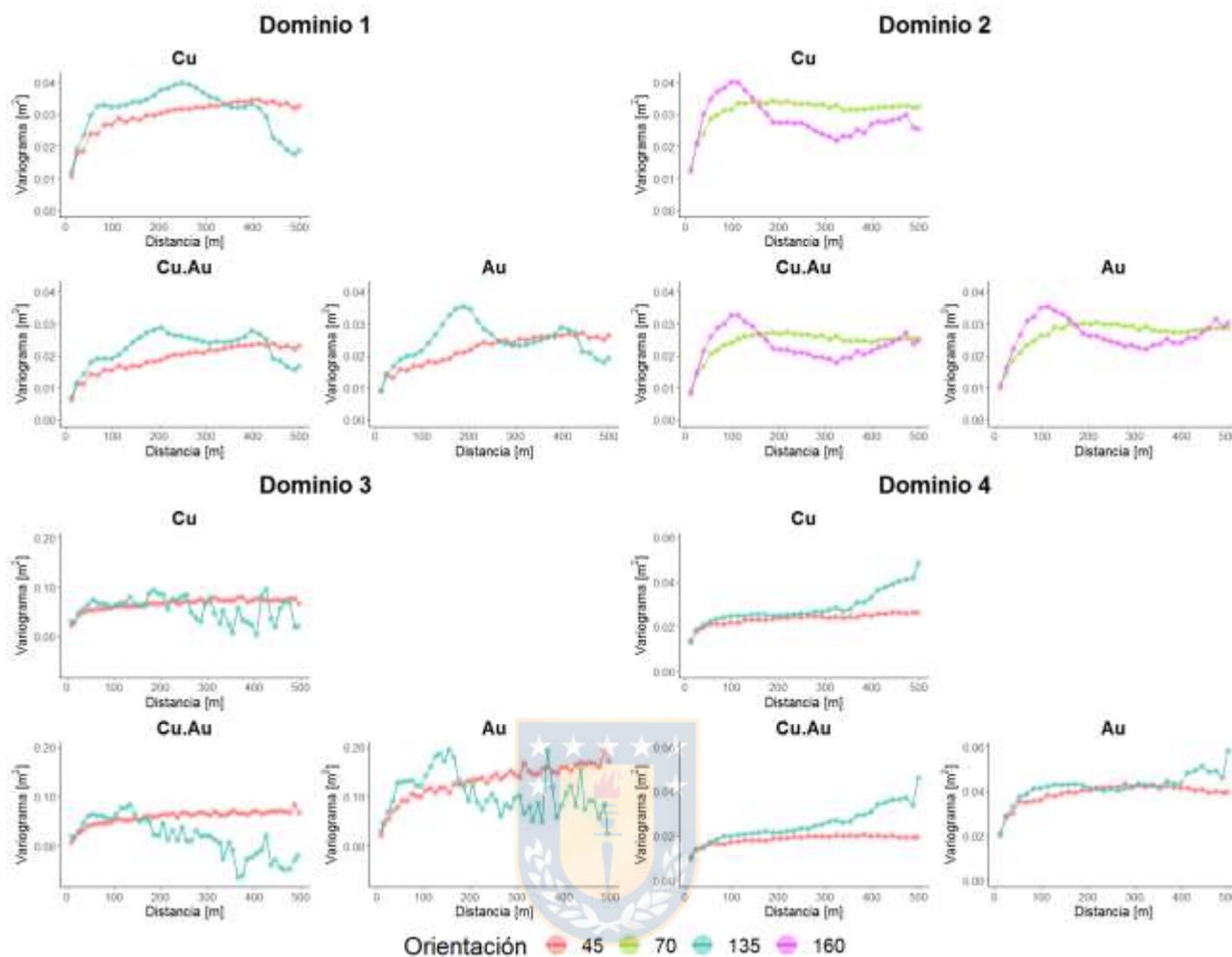


Figura 32. Variogramas direccionales de la base real según dominio.

Se debe tener en cuenta que al momento de realizar el ajuste de los modelos a los variogramas calculados, este no puede ser específico para cada una de las direcciones estudiadas al mismo tiempo, por lo que solo se proporciona un modelo para cada dominio. En los cuatro dominios se resuelve que el modelo que mejor se ajusta a los variogramas simples y cruzados en la dirección principal es el exponencial, cuyos parámetros como la pepa, meseta y rango son resumidos en la Tabla 30. Una vez establecido el ajuste de los modelos, se procede a determinar el número de vecinos máximos que se van a utilizar para las estimaciones en cada dominio, cuyos valores son calculados por prueba y error, los que corresponden a 80, 50, 60 y 100 para el dominio 1, 2, 3 y 4, respectivamente.

Con los modelos ajustados de los variogramas y el número óptimo de vecinos máximos, se realiza la estimación en la base de validación y en el modelo de bloques para cada dominio independientemente. En la Figura 33 se visualiza la ley de oro perteneciente al dominio 4,

específicamente en la cota 299.4, donde se muestra la base de validación original, la estimación de esta y, finalmente, la estimación realizada en el modelo de bloques cuyos puntos grises representan la base de entrenamiento. Es posible apreciar que para la base de validación estimada existen en menor cantidad leyes altas (color rojo) y leyes bajas (color azul) en comparación con la base original. Por otro lado, en los bloques estimados se advierte cómo la estimación continúa con la tendencia que siguen las leyes en la base de validación estimada, siendo más azules (valores pequeños) al Noroeste y más amarillos (valores medianos) al Sudeste.

Tabla 30. Parámetros de los modelos ajustados de variogramas simples y cruzados para la base real según dominio.

		Dominio 1	Dominio 2	Dominio 3	Dominio 4
Cobre	Pepa	0.008	0.001	0.013	0.008
	Meseta	0.027	0.035	0.056	0.016
	Rango	85	60	51	50
Oro	Pepa	0.006	0.001	0.001	0.010
	Meseta	0.019	0.030	0.133	0.031
	Rango	85	60	51	50
Cobre – Oro	Pepa	0.004	0.000	0.000	0.006
	Meseta	0.018	0.028	0.061	0.013
	Rango	85	60	51	50

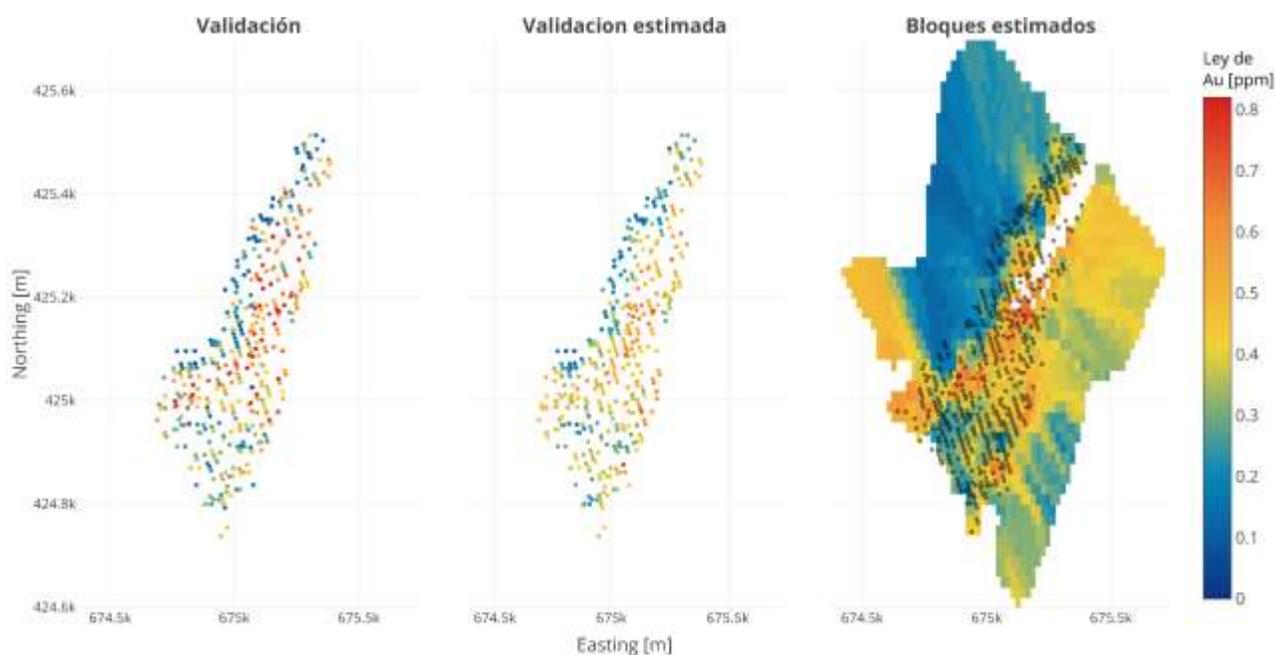


Figura 33. Estimación del oro con Cokriging Ordinario en la base real, dominio 4, cota 299.4.

4.3.3.2 Redes Neuronales Artificiales

Para las técnicas de ML se utiliza el mismo concepto planteado en el caso simulado, realizando dos tipos de estimaciones: una univariable y otra multivariable. En este caso, al contar solo con dos variables continuas, la elegida para ser estimada de manera univariable es el cobre puesto que presenta un comportamiento menos errático que el oro.

El algoritmo utilizado para utilizar ANN es el Rprop, donde la función de activación es la sigmoideal para los cuatro dominios estudiados. Por otro lado, tanto el número de capas como el número de neuronas depende del dominio a estimar, los que son calculados por prueba y error teniendo siempre en cuenta el tiempo de procesamiento como la minimización del error de estimación en la base de validación, llegando a mejores resultados utilizando entre 3 a 5 capas de hasta 60 neuronas, donde se busca un $threshold = \{0.05, 0.06, \dots, 0.15\}$ realizando más de 5000 combinaciones distintas en cada dominio. Los parámetros óptimos determinados según el mínimo error de estimación obtenido, son resumidos en la Tabla 31. Es interesante notar la diferencia entre los valores obtenidos para la base simulada y la base real, pasando de ocupar 3 capas a 5 y un número máximo de neuronas de 20 a 60. Ahora bien, también es importante recordar que el valor de los parámetros para cada una de las bases es adquirido según las combinaciones entre estos, las que son hechas en base al tiempo de procesamiento de los datos.

Tabla 31. Parámetros óptimos de ANN para la base real según dominio.

		Cobre				Oro			
		Dominio 1	Dominio 2	Dominio 3	Dominio 4	Dominio 1	Dominio 2	Dominio 3	Dominio 4
N° neuronas	Capa 1	18	12	62	20	40	42	30	7
	Capa 2	30	15	33	12	5	19	2	20
	Capa 3	11	9	6	2	2	1	20	20
	Capa 4	–	–	3	1	3	3	1	–
	Capa 5	–	–	–	–	2	1	1	–
Threshold		0.1	0.08	0.1	0.1	0.1	0.1	0.09	0.06

Con los parámetros anteriores se realiza la estimación en la base de validación y en el modelo de bloques creado en cada dominio. En la Figura 34 se visualiza la variable oro en el dominio 4 para la base de validación, la estimación de esta y, por último, la estimación del modelo de bloques. De la

base de validación estimada es notable como la suavización de los valores es menor que para el Cokriging Ordinario. En cuanto al modelo de bloques estimado, si bien no es posible determinar qué tan semejante es a la realidad, es posible mencionar que sigue la misma tendencia que el modelo de bloques estimado por Cokriging, donde la gran diferencia entre ellos es la continuidad que consigue ANN entre sus valores adquiridos, además de conseguir un rango más grande entre las leyes.

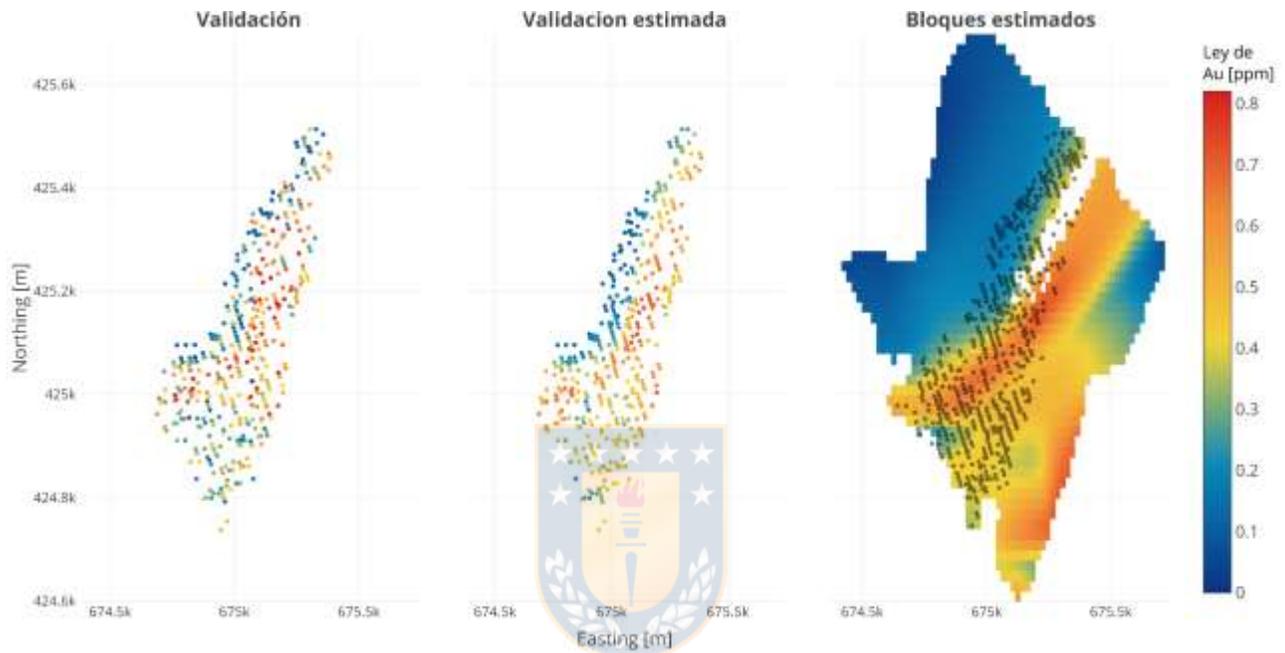


Figura 34. Estimación del oro con Redes Neuronales Artificiales en la base real, dominio 4, cota 299.4.

4.3.3.3 Árboles de Regresión

Como fue comentado anteriormente en la sección de Árboles de Regresión para el caso simulado, las estimaciones univariadas de las leyes minerales se hacen utilizando Kriging Ordinario pues después de varias pruebas es posible tener en conocimiento que CART arroja errores de estimación mayores en comparación con los otros métodos, por lo que utilizarlo en la primera etapa univariable podría conducir a una inexactitud superior.

Para la primera estimación se consideran los mapas variográficos y variogramas simples del cobre de la Figura 31 y Figura 32, respectivamente. Una vez determinado el modelo ajustado de los variogramas experimentales para cada dominio, se realiza la estimación donde se considera el mismo número máximo de vecinos según dominio utilizado para el Cokriging Ordinario (80, 50, 60 y 100).

Posteriormente, se busca la combinación óptima de parámetros según los considerados anteriormente: *minsplit*, *minbucket*, *maxdepth* y *cp*. La grilla creada cuenta con una extensión de valores donde *minsplit* y *minbucket* adquieren números entre $\{0,2, \dots, 60\}$, $maxdepth = \{1,2, \dots, 30\}$ y por último $cp = \{5 \times 10^{-9}, 5 \times 10^{-8}, \dots, 5 \times 10^{-2}\}$, formando más de 200 mil combinaciones para cada dominio. La combinación de parámetros que resultó tener un menor error de estimación para el oro se presenta en la Tabla 32, donde en comparación con las bases simuladas, los valores adquiridos son bastante superiores, a excepción del costo – complejidad que en el dominio 2 y 3 alcanzó valores inferiores.

Tabla 32. Parámetros óptimos de CART para la base real según dominio.

	Oro			
	Dominio 1	Dominio 2	Dominio 3	Dominio 4
<i>Minsplit</i>	28	1	50	17
<i>Minbucket</i>	3	5	8	6
<i>Maxdepth</i>	9	7	7	10
<i>Cp</i>	5×10^{-4}	5×10^{-8}	5×10^{-8}	5×10^{-4}

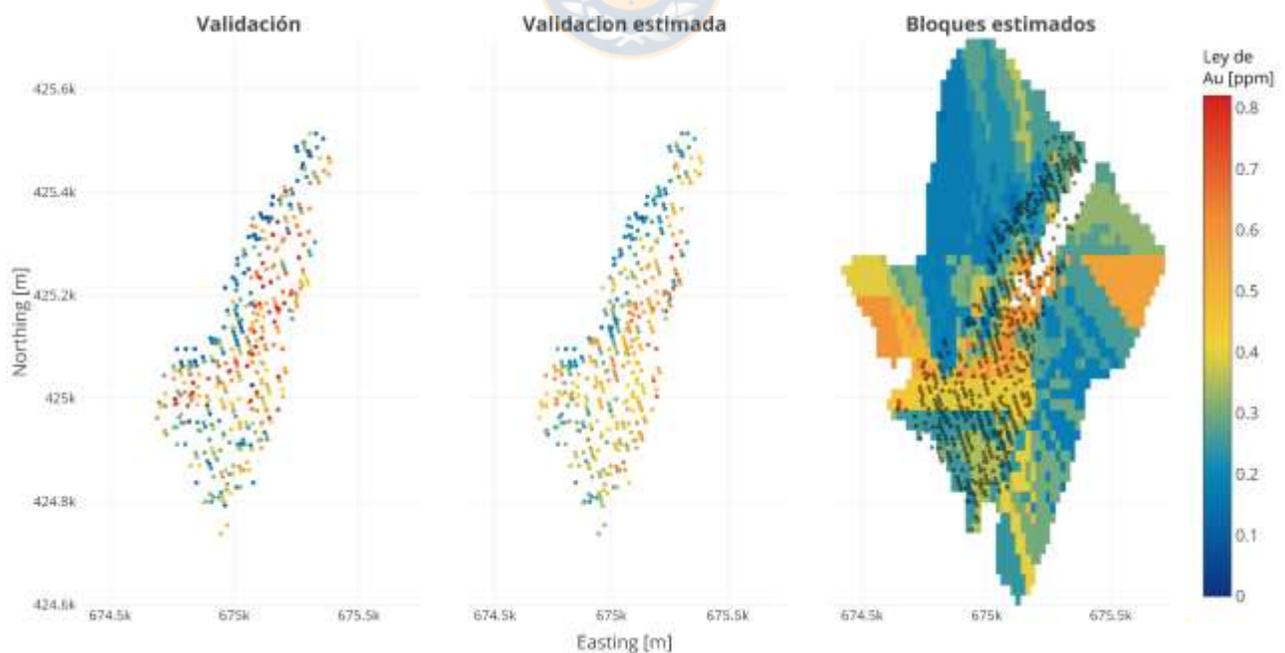


Figura 35. Estimación del oro con Árboles de Regresión en la base real, dominio 4, cota 299.4.

Una vez obtenido los parámetros óptimos, se realiza la estimación de las leyes en la base de validación y posteriormente en el modelo de bloques según cada dominio. Al igual que en los algoritmos anteriores, en la Figura 35 se visualiza la ley de oro en el dominio 4 de la base de validación, la estimación de esta y la estimación realizada en el modelo de bloques. En la base de validación estimada, la suavización de la ley es similar a la realizada por el Cokriging Ordinario. La gran diferencia radica en la estimación de la ley en los bloques, donde se distingue la particularidad de los árboles de regresión, que los resultados de la estimación tienden a segregarse por fronteras, siendo muy poco continuos los cambios entre sus valores. A pesar de esto, la tendencia en el modelo de bloque es similar a la de los otros algoritmos mencionados.

4.3.3.4 Support Vector Regression

El último algoritmo de estimación utilizado es SVR, donde nuevamente se utiliza el kernel Gaussiano, ecuación (73), pues es este el que obtiene mejor rendimiento. Los parámetros del método a ser optimizados son la constante de regularización C , épsilon de la función ϵ -insensible, gamma del kernel según la ecuación nombrada anteriormente y la tolerancia (tol). Se crea una grilla donde ϵ alcanza valores de 0 a 1, la constante C va desde 0.1 a 16, gamma entre 0 a 10 y por último la tolerancia con valores desde 0.001 a 1, realizando más de un millón de combinaciones para cada dominio con el fin de identificar la opción que obtenga un menor error de estimación. Los valores seleccionados para finalmente realizar las estimaciones son los contenidos en la Tabla 33, donde la gran diferencia con los parámetros utilizados en el caso simulado son los valores adquiridos por gamma pues para el caso real son bastante más grandes.

Tabla 33. Parámetros óptimos de SVR para el caso real según dominio.

	Cobre				Oro			
	Dominio 1	Dominio 2	Dominio 3	Dominio 4	Dominio 1	Dominio 2	Dominio 3	Dominio 4
C	1.208	1.041	3.509	0.514	1.888	1.627	1.894	2.567
ϵ	0.193	0.305	0.561	0.347	0.244	0.171	0.543	0.270
γ	10.5	8.0	3.480	6.5	1.282	2.25	0.840	1.578
Tol	0.2	0.5	0.37	0.3	1.0	0.75	0.6	0.4

Con los parámetros de la tabla anterior se realiza la estimación de las leyes en la base de validación y en modelo de bloques. Por último, en la Figura 36 se visualiza la ley de oro para el

dominio 4 en la base de validación, en la estimación de esta y, finalmente, la estimación realizada en el modelo de bloques. En comparación con las otras figuras, se aprecia que visualmente SVR es el que suaviza más las leyes en la estimación de la base de validación, obteniendo una variación menor entre estas, lo que también se observa en la estimación del modelo de bloques, donde en las ubicaciones en las que no hay datos de entrenamiento (puntos grises) las leyes tienden a adquirir un valor constante igual a la media de la base original hasta las fronteras del dominio.

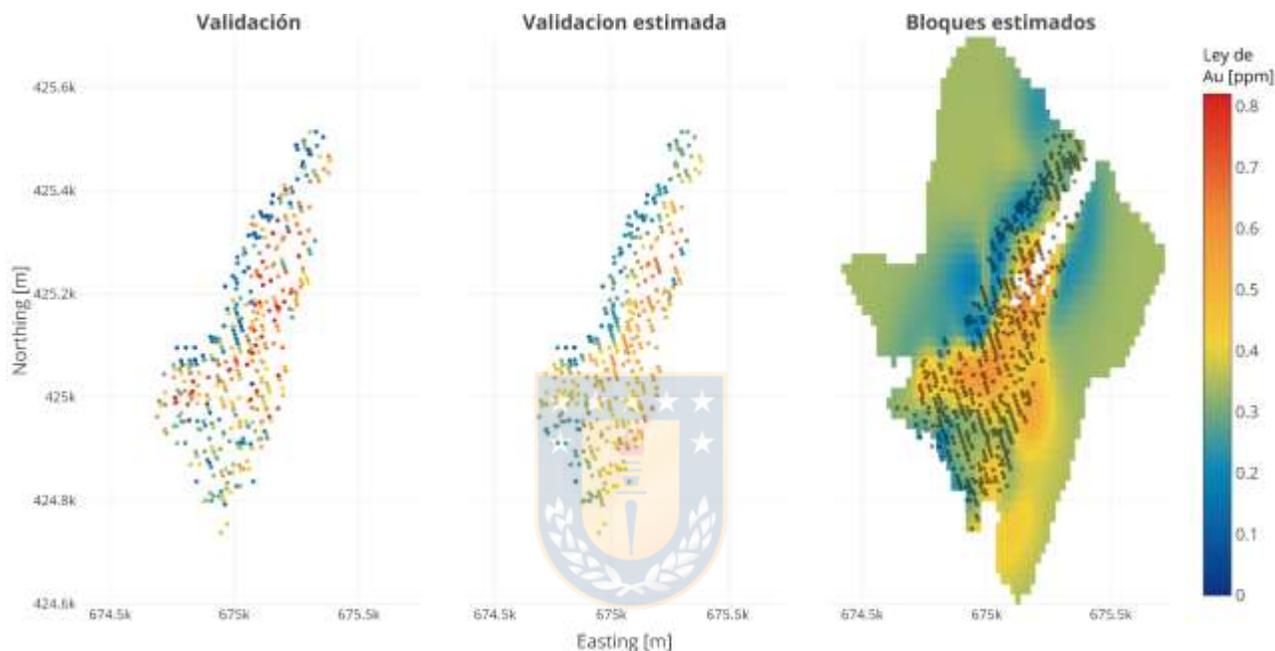


Figura 36. Estimación del oro con *Support Vector Regression* en la base real, dominio 4, cota 299.4.

4.3.4 MEDIDAS DE ERROR DE ESTIMACIÓN

Las medidas de error en la estimación consideradas son el ME, MAE, RMSE y R^2 definidas en las ecuaciones (74), (75), (76) y (77), respectivamente, donde estas son resumidas en una sola medida estadística denominada *skill value* pues no sería suficiente al considerar solamente uno de estos criterios para realizar la comparación entre los métodos de estimación. El *skill value* para el cobre y oro son determinados según la ecuación (78).

Los resultados obtenidos para cada dominio según el método de estimación se resumen en la Tabla 34, donde también es importante recordar que los valores obtenidos en CART para la variable cobre son en realidad del Kriging Ordinario pues pertenece a la estimación univariable. En este caso

se observa un comportamiento diferente para cada dominio, donde en el dominio 1 y 2 SVR no obtiene los mejores resultados cuando estima cobre, más bien estos son mejores utilizando OCOK, sin embargo, cuando se realiza la estimación multivariable para obtener la ley de oro en los mismos dominios, el *skill value* obtenido por SVR es relativamente mejor. Para el dominio 3 y 4 la situación es diferente, pues tanto para cobre como para oro los valores obtenidos en las medidas de rendimiento son mejores utilizando SVR. Por otro lado, los peores resultados obtenidos, a excepción del dominio 3, tanto para cobre como para oro son con ANN. Es importante recordar que el comportamiento de las leyes en el dominio 1 y 2 es diferente al dominio 3 y 4, pues en los dos primeros la correlación espacial es más fuerte mientras que en los dos últimos es más leve.

Tabla 34. Medidas de rendimiento en las estimaciones para el caso real según dominio.

		Cobre						Oro					
		ME	MAE	RMSE	R ²	<i>Skill value</i>	Ranking	ME	MAE	RMSE	R ²	<i>Skill value</i>	Ranking
Dominio 1	OCOK	-0.002	0.082	0.114	0.618	0.236	2 ^{do}	-0.002	0.079	0.113	0.489	0.244	2 ^{do}
	ANN	0.003	0.082	0.111	0.638	0.232	1 ^{er}	-0.041	0.089	0.122	0.515	0.299	4 ^{to}
	CART	-0.001	0.082	0.115	0.616	0.237	3 ^{er}	0.004	0.082	0.116	0.456	0.257	3 ^{er}
	SVR	-3x10 ⁻⁴	0.084	0.117	0.604	0.240	4 ^{to}	0.001	0.079	0.112	0.493	0.243	1 ^{er}
Dominio 2	OCOK	0.004	0.080	0.112	0.609	0.235	2 ^{do}	0.009	0.085	0.130	0.503	0.273	1 ^{er}
	ANN	0.018	0.093	0.129	0.506	0.289	4 ^{to}	-0.061	0.122	0.174	0.357	0.421	4 ^{to}
	CART	0.003	0.078	0.111	0.616	0.231	1 ^{er}	0.004	0.090	0.137	0.444	0.287	3 ^{er}
	SVR	-3x10 ⁻⁴	0.085	0.117	0.575	0.244	3 ^{er}	-2x10 ⁻⁴	0.089	0.134	0.466	0.277	2 ^{do}
Dominio 3	OCOK	-0.004	0.143	0.181	0.522	0.376	3 ^{er}	0.041	0.199	0.294	0.586	0.575	4 ^{to}
	ANN	5x10 ⁻⁵	0.136	0.177	0.543	0.359	2 ^{do}	9x10 ⁻⁴	0.207	0.291	0.586	0.540	2 ^{do}
	CART	-0.004	0.144	0.183	0.510	0.379	4 ^{to}	0.025	0.207	0.298	0.576	0.572	3 ^{er}
	SVR	9x10 ⁻⁴	0.129	0.171	0.571	0.344	1 ^{er}	0.003	0.203	0.284	0.614	0.529	1 ^{er}
Dominio 4	OCOK	-0.003	0.091	0.122	0.352	0.280	2 ^{do}	0.005	0.116	0.156	0.441	0.333	2 ^{do}
	ANN	-0.003	0.096	0.123	0.343	0.287	4 ^{to}	-0.008	0.129	0.169	0.396	0.367	4 ^{to}
	CART	-0.002	0.092	0.122	0.339	0.283	3 ^{er}	3x10 ⁻⁴	0.126	0.169	0.357	0.359	3 ^{er}
	SVR	1x10 ⁻⁴	0.089	0.118	0.382	0.269	1 ^{er}	1x10 ⁻⁴	0.118	0.157	0.428	0.332	1 ^{er}

Ahora bien, al juntar los cuatro dominios dentro de una misma base, y calcular las medidas de rendimiento de los algoritmos utilizados de manera global, contenido en la Tabla 35, se obtienen errores menores y coeficientes de determinación mayores en la estimación con SVR tanto para el

cobre como para el oro, resultando en un *skill value* menor y, por el contrario, los valores en los errores son mayores con un R^2 menor para el método ANN para ambas variables continuas.

Tabla 35. Medidas de rendimiento en las estimaciones para el caso real de manera global.

	Cobre				Oro			
	OCOK	ANN	OK	SVR	OCOK	ANN	CART	SVR
ME	-5×10^{-4}	0.005	-7×10^{-4}	2×10^{-4}	0.009	-0.032	0.005	7×10^{-4}
MAE	0.091	0.096	0.091	0.091	0.106	0.125	0.112	0.109
RMSE	0.125	0.129	0.126	0.124	0.161	0.178	0.168	0.161
R²	0.698	0.683	0.696	0.703	0.687	0.633	0.657	0.688
Skill value	0.247	0.262	0.248	0.245	0.307	0.371	0.320	0.301
Ranking	2^{do}	4^{to}	3^{er}	1^{er}	2^{do}	4^{to}	3^{er}	1^{er}

4.3.5 MEDIDAS ESTADÍSTICAS PRINCIPALES DEL MODELO DE BLOQUES

Si bien no es posible cuantificar el error producido al realizar las estimaciones en el modelo de bloques, resulta interesante comparar las medidas estadísticas principales de las leyes como la media, mediana, desviación estándar, mínimo y máximo para cada método utilizado en los diferentes dominios. La Tabla 36 contiene las estadísticas antes mencionadas para cada dominio, donde, a excepción del dominio 2 para la variable oro, el método SVR tiene una mayor ley media en comparación con el resto junto con una desviación estándar menor, en especial cuando se trata de que la variable estimada es el cobre. También es posible notar que para dicho algoritmo la mediana es casi siempre muy similar a la media, a diferencia de los otros métodos en los que en la mayoría de las ocasiones el comportamiento es diferente. Por otro lado, los valores máximos en las bases de datos estimadas más grandes numéricamente son obtenidos por OCOK y ANN en la mayoría de los dominios.

En cuanto a las medidas estadísticas para los cuatro dominios de manera global, contenidas en la Tabla 37, se obtienen leyes medias mayores para SVR en la variable cobre, tal como en la base categorizada por dominios, pero en el caso del oro se produce una ley media mayor utilizando ANN, esto probablemente influenciado por el dominio 2, ya que tiene una media bastante superior a las demás. Por otro lado, existen valores máximos mayores en SVR para el cobre y OCOK para el oro,

ambas contenidas en el dominio 3. Por último, la desviación estándar sigue siendo menor con SVR para el caso del cobre, mientras que para el oro es menor utilizando CART.

Tabla 36. Estadísticas principales de las leyes estimadas en el modelo de bloques según dominio.

		Cobre				Oro			
		OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
Dominio 1	Media	0.213	0.227	0.213	0.264	0.184	0.210	0.173	0.213
	Mediana	0.187	0.206	0.187	0.265	0.148	0.160	0.150	0.187
	σ	0.130	0.169	0.130	0.049	0.125	0.163	0.109	0.130
	Mínimo	0.008	0.000	0.006	0.056	0.000	0.049	0.000	0.008
	Máximo	0.716	0.800	0.719	0.573	0.732	0.726	0.695	0.716
Dominio 2	Media	0.244	0.220	0.245	0.245	0.195	0.346	0.210	0.244
	Mediana	0.217	0.206	0.219	0.244	0.183	0.310	0.194	0.217
	σ	0.118	0.142	0.118	0.033	0.117	0.238	0.127	0.118
	Mínimo	0.018	0.000	0.018	0.082	0.003	0.026	0.025	0.018
	Máximo	0.687	0.662	0.704	0.532	0.753	0.754	0.753	0.687
Dominio 3	Media	0.705	0.609	0.704	0.711	0.656	0.598	0.608	0.705
	Mediana	0.712	0.633	0.708	0.744	0.653	0.602	0.597	0.712
	σ	0.139	0.184	0.138	0.122	0.257	0.297	0.210	0.139
	Mínimo	0.225	0.244	0.239	0.180	0.079	0.141	0.149	0.225
	Máximo	1.215	1.080	1.215	1.304	1.796	1.145	1.409	1.215
Dominio 4	Media	0.339	0.355	0.336	0.363	0.322	0.313	0.308	0.339
	Mediana	0.340	0.342	0.332	0.366	0.322	0.292	0.276	0.340
	σ	0.107	0.122	0.104	0.042	0.125	0.180	0.140	0.107
	Mínimo	0.090	0.017	0.095	0.163	0.065	0.000	0.069	0.090
	Máximo	0.741	0.786	0.703	0.628	0.816	0.769	0.771	0.741

Tabla 37. Estadísticas principales de las leyes estimadas en el modelo de bloques de manera global.

	Cobre				Oro			
	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
Media	0.325	0.309	0.324	0.345	0.288	0.340	0.281	0.325
Mediana	0.275	0.281	0.274	0.265	0.235	0.279	0.224	0.275
σ	0.206	0.205	0.205	0.171	0.221	0.250	0.202	0.206
Mínimo	0.008	0.000	0.006	0.056	0.000	0.000	0.000	0.008
Máximo	1.215	1.080	1.215	1.304	1.796	1.145	1.409	1.215

4.3.6 CATEGORIZACIÓN DE RECURSOS MINERALES

La metodología utilizada para realizar la categorización de recursos minerales es la anteriormente planteada en la sección 2.5.1, donde se utiliza la varianza de cokriging del cobre pues es este el que se asume como mineral de interés. A diferencia del caso simulado, no se cuenta con los datos reales en el modelo de bloques, por lo que no es posible realizar una comparación en la calidad de los datos estimados. Sin embargo, con fines investigativos, se decide realizar la categorización en la base de validación para ver la calidad de los datos pues, de esta manera, se puede hacer la comparación con lo obtenido en las bases simuladas.

Un dato es clasificado como recurso mineral medido, indicado o inferido si es que se cumplen las condiciones presentadas en las ecuaciones (83), (84) y (85), respectivamente. Una vez realizada la categorización, esta es replicada en los métodos de ML para posteriormente hacer la evaluación económica en el modelo de bloques y ver la calidad de los datos en la base de validación.

$$\frac{\sigma(x_0)}{Z_1^*(x_0)} < 0.2376 \quad (83)$$

$$0.2376 < \frac{\sigma(x_0)}{Z_1^*(x_0)} < 0.7628 \quad (84)$$

$$\frac{\sigma(x_0)}{Z_1^*(x_0)} > 0.7628 \quad (85)$$

En la Figura 37 se encuentra la visualización de la categorización de recursos minerales resultante para la base de validación y para el modelo de bloques. En ella se puede observar que para ambas bases los recursos medidos son mínimos en comparación con el resto, a excepción del dominio 3 ya que en este los recursos inferidos llegan a ser nulos para la base de validación. Además, para la base de validación siempre se tienen más recursos indicados, a diferencia del modelo de bloques donde la situación cambia para el dominio 1 y 2 pues se tiene una mayor cantidad de recursos inferidos.

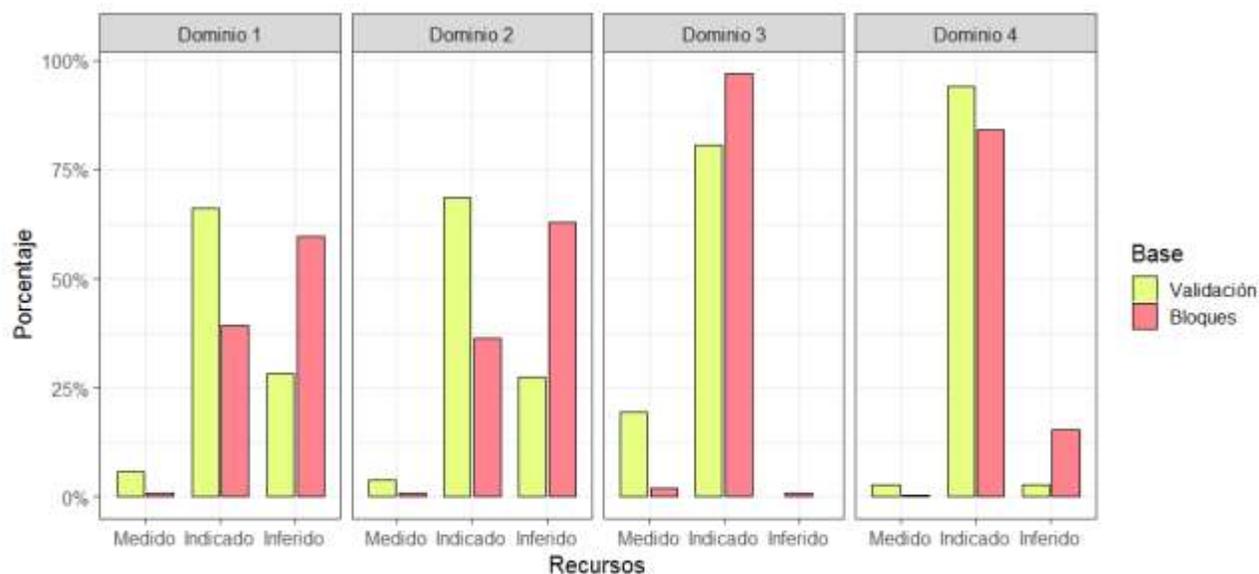


Figura 37. Categorización de recursos minerales para la estimación de la base real de validación y modelo de bloques.

En cuanto a la calidad de los datos, esta se realiza en base a la metodología planteada en el capítulo del caso simulado, donde como principales medidas se utilizan la media y la desviación estándar (σ) de las leyes. Con el fin de sintetizar la información obtenida para cada dominio, la Tabla 38 contiene las medidas anteriormente mencionadas de manera global para la base de validación y para el modelo de bloques, donde esta última es proporcionada para realizar algunas observaciones entre las estadísticas obtenidas. De ambas bases se advierte una mayor ley media para los recursos minerales medidos (med) y una menor ley media para los inferidos (inf), mientras que los recursos indicados (ind) se encuentran entre ambos. De la base de validación estimada es posible notar que nuevamente la desviación estándar de la base original para cada recurso mineral es mayor al de las bases estimadas, esto debido al suavizamiento que provoca cada método al estimar las leyes. También es destacable que las leyes medias en los recursos inferidos en las bases estimadas son mayores a la ley media de la base original para ambas variables continuas, dando a entender que los métodos son más arriesgados en esta categoría. En particular, redes neuronales obtiene una ley media mayor en las tres clasificaciones de recursos para la variable oro, además de una desviación estándar mayor a los otros métodos para ambas variables continuas, siendo este último comportamiento replicado en el modelo de bloques estimado. Además, en el modelo de bloques, en lo recursos inferidos, se replica la tendencia a tener mayor ley media con SVR para el cobre y ANN para el oro.

Tabla 38. Calidad de los datos para la estimación de la base real según recursos minerales.

			Cobre					Oro				
			Original	OCOK	ANN	CART	SVR	Original	OCOK	ANN	CART	SVR
Base de validación	Med	Media	0.689	0.708	0.641	0.704	0.653	0.666	0.657	0.666	0.623	0.606
		σ	0.258	0.207	0.232	0.209	0.212	0.424	0.341	0.314	0.301	0.314
	Ind	Media	0.378	0.376	0.374	0.377	0.376	0.351	0.340	0.386	0.347	0.351
		σ	0.198	0.149	0.168	0.148	0.157	0.266	0.198	0.225	0.205	0.211
	Inf	Media	0.109	0.112	0.114	0.112	0.130	0.091	0.093	0.116	0.092	0.108
		σ	0.075	0.041	0.070	0.042	0.056	0.075	0.046	0.075	0.040	0.056
Modelo de bloques	Med	Media	–	0.668	0.474	0.666	0.580	–	0.600	0.493	0.544	0.534
		σ	–	0.198	0.292	0.201	0.229	–	0.314	0.368	0.245	0.321
	Ind	Media	–	0.446	0.393	0.444	0.411	–	0.404	0.398	0.387	0.395
		σ	–	0.182	0.204	0.182	0.193	–	0.216	0.252	0.196	0.221
	Inf	Media	–	0.156	0.193	0.157	0.252	–	0.127	0.260	0.134	0.238
		σ	–	0.063	0.134	0.064	0.045	–	0.071	0.221	0.076	0.055

Al realizar el ranking según la metodología planteada, mostrado en la Tabla 39, resulta que SVR es más conservador en los recursos minerales medidos, OCOK en los indicados y CART en los inferidos, mientras que el menos conservador es distinto según la variable continua, donde para el cobre es más arriesgado CART en los medidos y SVR en los indicados e inferidos y para el oro es ANN, independiente de la categoría. Es importante notar que lo anterior parece coincidir con los resultados obtenidos en el caso simulado.

Tabla 39. Ranking de la calidad de los datos para la estimación de la base real según recursos minerales.

	Cobre				Oro			
	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
Medidos	3	2	4 ^{to}	1 ^{er}	3	4 ^{to}	2	1 ^{er}
Indicados	1 ^{er}	2.5	2.5	4 ^{to}	1 ^{er}	4 ^{to}	2	3
Inferidos	2	3	1 ^{er}	4 ^{to}	2	4 ^{to}	1 ^{er}	3

4.3.7 ANÁLISIS ECONÓMICO

Para realizar el análisis económico se utiliza la ley equivalente, Anexo C. 1, siendo el cobre el mineral de interés a ser extraído y el oro un subproducto. Los datos considerados para hacer el cálculo de esta se encuentran en la Tabla 40, cuyos precios de cobre y oro son los mismos considerados en el

caso simulado. Utilizando las ecuaciones (113) y (114) se determina que el cálculo de la ley equivalente se realiza según la siguiente expresión:

$$Ley_{eq} = Ley_{cobre} + 1.1532 \times Ley_{oro} \quad (86)$$

Tabla 40. Datos económicos considerados para el caso real.

	Cobre	Oro
Precio	4,617.5 [USD\$/ton]	1,554.2 [USD\$/tr oz]
Costo venta [US\$/ton]	700	8,000,000
% pagable de mineral	97	96
Recuperación metalúrgica [%]	92	85

En la Tabla 41 se muestra la ley media y el tonelaje obtenido luego de hacer la transformación a ley equivalente en los cuatro dominios de manera global, donde se considera una densidad de 2.7 [ton/m³] para realizar este último cálculo. Como fue comentado en la sección anterior, SVR obtendría leyes mayores para los recursos minerales inferidos, por lo que esta conducta se replicó al utilizar la ley equivalente, obteniendo un valor bastante superior a los demás métodos. Además, el método mencionado presenta una mayor ley media sin categorizar por recursos. Cabe mencionar que al realizar la categorización basada en el método de Cokriging y replicar este en los demás, es que todos comienzan con un mismo tonelaje, teniendo una mayor cantidad en los indicados, luego en los inferidos y, por último, un valor bastante inferior en los medidos, obteniendo un total de 559.3 [Mton].

Tabla 41. Ley media y tonelaje utilizando ley equivalente para estimación de bloques de manera global.

	Ley media [%]				Tonelaje [Mton]			
	Total	Medidos	Indicados	Inferidos	Total	Medidos	Indicados	Inferidos
OCOK	0.657	1.359	0.912	0.303	559.3	5.78	315.6	237.9
ANN	0.701	1.043	0.852	0.493	559.3	5.78	315.6	237.9
CART	0.648	1.293	0.890	0.312	559.3	5.78	315.6	237.9
SVR	0.726	1.195	0.867	0.527	559.3	5.78	315.6	237.9

Para calcular la ley de corte (COG) se considera un costo de mina de 2.7 [USD\$/ton] y un costo de procesamiento de 13 [USD\$/ton], obteniendo un COG de 0.452%. La Tabla 42 contiene la ley media y el tonelaje después de filtrar los modelos de bloques estimados según la ley de corte. Debido a las altas leyes equivalentes presentes en los recursos minerales inferidos del método SVR, como

señala la Tabla 41, es que, al aplicar la ley de corte, la ley media en esta categoría no aumenta significativamente como en los otros algoritmos, implicando también una menor disminución en el tonelaje de este. Por otro lado, la ley media del estéril es bastante superior para SVR, pero, en contraste con los demás métodos, se tiene poca cantidad de este. En general, casi no hubo estéril en los medidos, mientras que los inferidos son los que disminuyeron en mayor porcentaje su tonelaje, a excepción del método SVR. Como último comentario, es posible advertir que la ley media total aumentó bastante en todos los métodos menos en SVR, cuyo aumento fue mucho menor.

Tabla 42. Ley media y tonelaje después de filtrar con ley de corte para estimación de bloques de manera global.

COG = 0.452%	Ley media [%]					Tonelaje [Mton]				
	Estéril	Total	Medidos	Indicados	Inferidos	Estéril	Total	Medidos	Indicados	Inferidos
OCOK	0.275	0.910	1.359	0.946	0.531	222.7	336.6	5.78	296.1	34.71
ANN	0.241	0.950	1.250	1.000	0.820	195.7	363.6	4.56	250.4	108.6
CART	0.274	0.884	1.293	0.926	0.540	216.2	343.1	5.78	294.5	42.8
SVR	0.368	0.755	1.196	0.888	0.550	42.00	517.3	5.77	302.4	209.1

La Tabla 43 contiene los ingresos, costos y beneficios obtenidos en los recursos minerales económicamente explotables, que son los medidos e indicados pues poseen una mayor confianza. Al igual que en el caso anterior, se considera que los bloques de mineral pueden ser simultáneamente extraíbles sin la necesidad de remover estéril. El método ANN es el que obtiene ingresos y costos menores, sin embargo, SVR es el que presenta un menor beneficio económico debido a sus altos costos, mientras que, por el contrario, el método OCOK consigue un mayor beneficio con 538 [MUSD\$] adicionales, siendo aproximadamente un 10% más que SVR. En el Anexo E. 1 se encuentra la evaluación económica obtenida en la estimación de bloques en cada dominio.

Tabla 43. Evaluación económica para estimación de bloques de manera global.

	Ingreso [MUSD\$]			Costo [MUSD\$]			Beneficio [MUSD\$]		
	Total	Medidos	Indicados	Total	Medidos	Indicados	Total	Medidos	Indicados
OCOK	10,012	273.1	9,739	4,739	90.71	4,649	5,272	182.4	5,090
ANN	8,907	198.1	8,709	4,003	71.55	3,932	4,904	126.5	4,777
CART	9,740	259.8	9,481	4,714	90.71	4,623	5,026	169.1	4,857
SVR	9,573	239.9	9,333	4,839	90.55	4,748	4,734	149.3	4,585

4.4 ANÁLISIS DE RESULTADOS

Para realizar el análisis de resultados se debe tener presente que cada dominio se estimó de manera independiente pues al realizar el procesamiento de la base de datos completa, el tiempo de estimación resultaba ser muy alto siendo contraproducente. Teniendo esto en cuenta, es que se analizan los resultados obtenidos para cada dominio. Refiriéndose al *skill value*, los métodos que obtuvieron mejores resultados según el dominio se resumen en la Tabla 44, donde para el dominio 1, 3 y 4 se consiguieron mejores resultados para los métodos de *Machine Learning*. Ocurre lo contrario para el dominio 2, donde el *skill value* es mejor utilizando OK en el cobre (estimación univariable para posteriormente ser usada en CART) y OCOK en el caso de oro (estimación multivariable). Ahora bien, al determinar las medidas de error de estimación para la base de manera global, es decir, uniendo las cuatro bases en una, resultan mejores *skill value* utilizando SVR, tanto para el cobre como para el oro.

Tabla 44. Mejor *skill value* para la base real según dominio.

	Dominio 1	Dominio 2	Dominio 3	Dominio 4	Base total
Cobre	ANN	OK	SVR	SVR	SVR
Oro	SVR	OCOK	SVR	SVR	SVR

Si bien se obtuvo mejores resultados con un algoritmo de ML en el dominio 1, la mejora es menor en comparación con el aumento que obtiene el método tradicional en los algoritmos restantes, tal como se muestra en la Tabla 45. En cuanto al dominio 2, el porcentaje de mejora del OCOK en comparación con ANN es bastante grande para ambas variables continuas, mientras que para el SVR es alrededor de un 3%. Para estos dos casos mencionados, dominio 1 y 2, es presumible que el método tradicional obtiene mejores resultados por el comportamiento que se logra percibir en los mapas variográficos y variogramas direccionales estudiados, Figura 31 y Figura 32 respectivamente, pues es visible una alta correlación espacial, a diferencia del dominio 3 y 4, en los cuales SVR obtuvo mejor rendimiento que los otros algoritmos, donde la meseta es cercana a la pepa, presumiendo una correlación espacial menor.

Particularmente en el dominio 3, que es el que presenta un comportamiento más errático con altas leyes y elevada desviación estándar, todos los métodos de ML presentaron mejores resultados,

donde ANN es mejor en aproximadamente un 5.3% y SVR en un 8.2%, ambos porcentajes refiriéndose a las dos variables continuas presentes, y CART, que fue utilizando en la estimación multivariable del oro, es mejor en un 0.57%. Por otro lado, en el dominio 4 la mejora con SVR es alrededor de 4% para el cobre y 0.10% para el oro, siendo este último un porcentaje bastante menor en contraste con el 8.9% de mejora que presenta el método tradicional versus ANN y CART. Si bien en ambos dominios la meseta es bastante cercana a la pepa en comparación con el dominio 1 y 2, el comportamiento del dominio 3 es mucho más errático, asemejándose más a las condiciones que estipula la hipótesis de este trabajo de investigación.

Tabla 45. Porcentaje de mejoramiento y empeoramiento de ML versus Cokriging Ordinario.

	Cobre [%]			Oro [%]		
	ANN – OCOK	OK – OCOK	SVR – OCOK	ANN – OCOK	CART – OCOK	SVR – OCOK
Dominio 1	↑ 1.74	↓ 0.23	↓ 1.67	↓ 22.6	↓ 5.07	↑ 0.58
Dominio 2	↓ 23.1	↑ 1.56	↓ 4.02	↓ 54.0	↓ 4.87	↓ 1.48
Dominio 3	↑ 4.52	↓ 0.99	↑ 8.38	↑ 6.07	↑ 0.57	↑ 8.06
Dominio 4	↓ 2.39	↓ 1.00	↑ 3.94	↓ 10.1	↓ 7.72	↑ 0.10
Base total	↓ 5.83	↓ 0.35	↑ 0.76	↓ 20.7	↓ 4.29	↑ 1.95

En la Tabla 46 se muestra el ranking que establece que método es más conservador sin diversificar por recursos minerales, donde se tiene como resultado que el método más conservador para el cobre es SVR y para el oro es OCOK. Esto coincide con lo planteado en el caso simulado, pues la mayoría de los casos resultaron ser más conservadores cuando se utilizaba SVR y, por otro lado, OCOK mostró una tendencia a ser menos arriesgado a medida que la variable se volvía más errática. Además, el método menos conservador en la variable cobre es OCOK y, como era de esperarse según lo planteado en la Tabla 39, en la variable oro es ANN, ambos resultados similares a lo planteado en la Tabla 22 del caso simulado.

Tabla 46. Ranking de la calidad de los datos para la estimación de la base real en general.

	OCOK	ANN	CART	SVR
Cobre	4 ^{to}	2	3	1 ^{er}
Oro	1 ^{er}	4 ^{to}	3	2

Que SVR sea el más conservador en el caso del cobre, y el segundo más conservador en el caso del oro, puede ser la razón por la cual en la evaluación económica de este consigue el menor beneficio,

pues, además del dominio 3, las leyes no presentan grandes variaciones entre sus valores, como lo muestra la Tabla 36, demostrando ser menos arriesgado al estimar en ubicaciones lejanas de los puntos utilizados para entrenar el algoritmo.

La Figura 38 muestra el caso particular del dominio 2, que es en el que ML obtiene peores resultados, visualizando la estimación realizada por los cuatro algoritmos empleados en el modelo de bloques a diferentes cotas, donde los puntos grises representan los datos de la base de entrenamiento cercanos a la cota. En este caso, CART se asemeja bastante a los valores de las leyes que adquiere OCOK, mientras que ANN es el método que estima leyes altas en mayor cantidad (color rojo), pero que sin embargo sigue un comportamiento similar a los demás. Ahora, el método que llama más la atención es SVR, pues este no se arriesga como los demás a estimar diversos valores en ubicaciones en las que no se tienen datos cercanos de la base de entrenamiento, más bien tiende a estimar un valor uniforme, manteniéndose dentro de la media, lo que podría significar una desventaja frente a una planificación a largo plazo en la extracción del mineral.

Por último, si bien no está dentro de los objetivos del estudio, es importante mencionar que realizar la estimación por sectores conllevó a un tiempo de procesamiento mucho menor que hacerlo con la base total, permitiendo también realizar una mejor optimización en cuando a los parámetros que cada método utiliza.

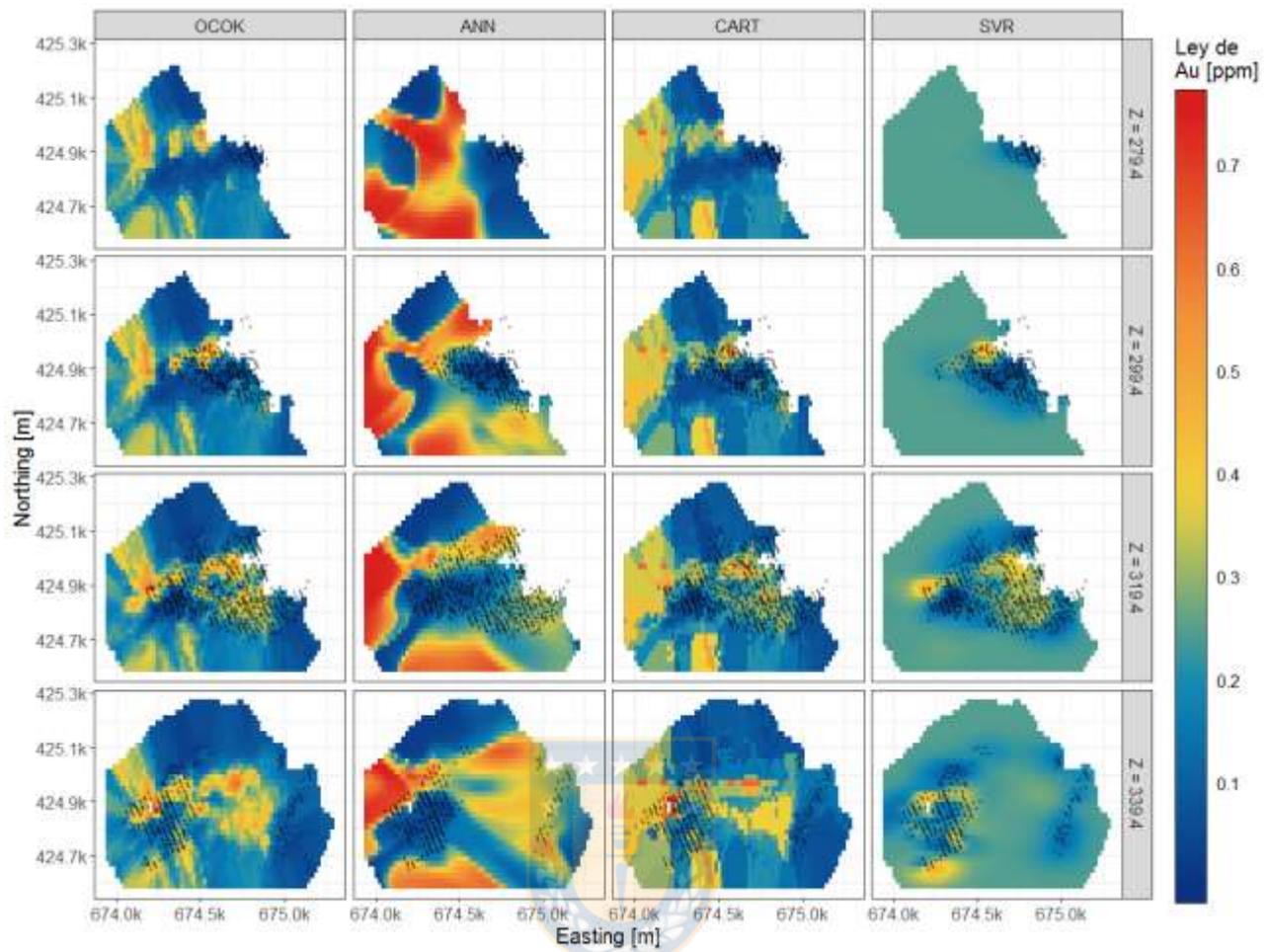


Figura 38. Estimación de dominio 2 en el modelo de bloques a diferentes cotas.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

El trabajo de investigación fue realizado bajo ciertas condiciones de complejidad que suelen presentarse en diversos yacimientos, en especial lo que poseen variables de comportamiento errático, tales como una baja o nula correlación espacial, escasa información disponible y, al existir más de una variable, como sucede en los yacimientos polimetálicos, que exista una alta correlación entre ellas. Se sentaron dichas bases con el fin de establecer si se obtienen mejores resultados en la estimación de leyes minerales, medidos según la disminución en los errores producidos en esta, utilizando técnicas de Inteligencia Computacional, como Redes Neuronales Artificiales, Árboles de Regresión, y *Support Vector Regression*, en vez del método convencional de Cokriging.

Se estudiaron dos casos, uno ficticio y uno real, donde en ambos se contaba con al menos dos variables continuas para realizar las estimaciones, además de las coordenadas *Easting*, *Northing* y *Elevation*. Las estimaciones fueron divididas principalmente en dos partes: la primera estimación se realizó de manera univariable, donde las variables de entrada eran las coordenadas y la variable de salida era la ley de un mineral, mientras que la segunda estimación se realizó de manera multivariable, siendo variables de entradas las coordenadas y el resultado de la primera estimación, y la variable de salida correspondía a la ley de otro mineral.

Basado en el análisis de resultados de ambos casos, la principal conclusión es que efectivamente se obtiene un mejor rendimiento en la estimación de recursos minerales utilizando dos de las tres técnicas empleadas de *Machine Learning* en contraste con el Cokriging Ordinario, dentro de las que destaca *Support Vector Regression*.

En base a los resultados obtenidos, las siguientes conclusiones fueron hechas:

- De los cien casos simulados y los cuatro dominios del caso real, *Support Vector Regression* es el que obtiene mejores resultados en cuanto al error de estimación mientras que el método que presenta un rendimiento inferior al resto es Árboles de Regresión.
- Cuando *Support Vector Regression* obtiene mejor rendimiento que Cokriging Ordinario, lo que ocurre el 95% de las veces, la disminución en el error de estimación es más grande a diferencia de

cuando Cokriging obtiene mejores resultados que este, siendo una mejora del 7% en los casos simulados y 8% en el caso real en el dominio 3, el cual replica las condiciones indicadas en la hipótesis.

- En general, los métodos *Support Vector Regression* y Árboles de Regresión parecen ser los más conservadores a la hora de estimar leyes minerales, manteniéndose más cercanos a la ley media, mientras que Redes Neuronales Artificiales y Cokriging Ordinario tienden a estimar leyes altas en una cantidad mayor, siendo más arriesgados que los dos primeros.
- En cuanto a la evaluación económica realizada en el caso simulado, *Support Vector Regression* genera una menor incertidumbre en términos monetarios, siendo el método que más se acerca al valor económico obtenido realmente en una mayor cantidad de veces, mientras que Árboles de Regresión es el que genera una mayor incertidumbre, generando diferencias mayores en cuanto al valor real.
- Cuando la variable regionalizada presenta anisotropía, se presume que los métodos de *Machine Learning* utilizados son desventajosos en comparación con utilizar las técnicas geoestadísticas tradicionales, pues son incapaces de representar las direcciones preferenciales de continuidad en el espacio.
- Si bien el algoritmo de *Support Vector Regression* parece ser superior a los demás en términos de disminución en el error de estimación, este resulta ser excesivamente conservador al estimar en sitios en los que no se cuenta con información cercana perteneciente a la base de entrenamiento, tendiendo a estimar valores similares a la media, lo que puede significar una gran desventaja en términos de planificación minera a largo plazo, donde se especifican los periodos en los que se va a extraer el mineral.
- Por último, y no menos importante, el método que tiene un tiempo de procesamiento de datos menor resultó ser Árboles de Regresión seguido de *Support Vector Regression*, siendo el caso contrario el método de Redes Neuronales Artificiales, pues no solo depende de la cantidad de datos con las que se entrena el modelo, sino también del número de capas y la cantidad de neuronas presentes en cada una de ellas.

Es necesario tener en cuenta que todas las combinaciones de parámetros ajustados para cada método de estimación fueron hechas según la disponibilidad computacional, debido a esto para Redes Neuronales se realizaron una menor cantidad de pruebas ya que el tiempo de procesamiento de los datos era bastante superior a los demás, a contrario del método CART que tiene un costo

computacional mucho menor. Por otro lado, también es importante mencionar que, si bien se hicieron numerosas combinaciones entre diferentes parámetros elegidos según el método de estimación empleado, existen otra serie de parámetros que no fueron considerados en las combinaciones ya que incrementaría el tiempo de estimación que de por sí ya era elevado. Teniendo esto en cuenta, es que es posible suponer una mejora en el rendimiento de los algoritmos de estimación empleados tanto los de ML como los tradicionales, al incrementar el número de combinaciones e incorporar nuevos parámetros que anteriormente fueron considerados constantes.

Si bien la hipótesis resulta ser asertiva, es fundamental realizar otras series de pruebas con las consideraciones previamente indicadas para poder proponerla como metodología de estimación a ser aplicada en situaciones que repliquen las condiciones señaladas en un inicio.

Algunas recomendaciones y trabajos a futuros propuestos son los siguientes:

- Incorporar parámetros de los algoritmos empleados que no fueron considerados para realizar la estimación y, además, incluir una gama más diversa en los valores de estos, para finalmente cuantificar la diferencia producida entre los errores y compararlos con los resultados obtenidos de en este trabajo.
- Incorporar otros algoritmos pertenecientes a la Inteligencia Computacional o variantes de los ya utilizados, además de realizar combinaciones entre estos incluyendo las técnicas geoestadísticas.
- Realizar un nuevo estudio de caso ficticio donde las variables regionalizadas presentes en el cuerpo mineralizado tengan comportamiento anisotrópico con el fin de medir el impacto generado en los algoritmos de ML en contraste con los métodos tradicionales.

REFERENCIAS

- [1] HARTMAN, Howard L. y MUTMANSKY, Jan M. *Introductory mining engineering*. 2nd ed. Hoboken: John Wiley & Sons, 2002.
- [2] CIM Standing Committee on Reserve Definitions. *CIM Definition Standards – For Mineral Resources and Mineral Reserves*. 2014.
- [3] CIM Estimation Best Practices Committee. *Estimation of Mineral Resources & Mineral Reserves Best Practice Guidelines*. 2003.
- [4] CANET MIQUEL, Carles y CAMPRUBÍ I CANO, Antoni. *Yacimientos minerales: Los tesoros de la tierra*. México: Fondo de Cultura Económica, 2006.
- [5] EMERY, Xavier. *Geoestadística*. Santiago de Chile: Departamento de Ingeniería de Minas. Universidad de Chile, 2013.
- [6] ALPAYDIN, Ethem. *Introduction to machine learning*. 2^{da} ed. MIT press, 2010.
- [7] MORALES, Eduardo y GONZÁLEZ, Jesús. *Aprendizaje computacional*. 2010.
- [8] DUTTA, Sridhar, et al. Machine Learning Algorithms and Their Application to Ore Reserve Estimation of Sparse and Imprecise Data. *Journal of Intelligent Learning Systems and Applications*, 2010, vol. 2, no 2, p. 86-96.
- [9] DUTTA, Sridhar, et al. Investigation of two Neural Network Ensemble Methods for the Prediction of Bauxite Ore Deposit. *Proceedings of the 6th International Conference on Information Technology*, 2003.
- [10] DUTTA, Sridhar, et al. Support Vector Machines – An Emerging Technique for Ore Reserve Estimation. *Proceedings of the Sixth International Symposium on Information Technology Applied to Mining (CD)*, Peruvian Institute of Mining Engineers, 2006.
- [11] SAMANTA, Biswajit. Radial basis function network for ore grade estimation. *Natural resources research*, 2010, vol. 19, no 2, p. 91-102.
- [12] KRIGE, Danie G. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 1951, vol. 52, no 6, p. 119-139.
- [13] MATHERON, Georges. Principles of geostatistics. *Economic geology*, 1963, vol. 58, no 8, p. 1246-1266.

- [14] GIRALDO HENAO, Ramón. Introducción a la Geoestadística: Teoría y aplicación. *Bogotá: Departamento de estadística. Universidad Nacional de Colombia*, 2002.
- [15] EMERY, Xavier y PIZARRO, Sebastián . Cokriging. *Santiago de Chile: Departamento de Ingeniería de Minas. Universidad de Chile*, 2013.
- [16] PAPRITZ, A.; KÜNSCH, H. R. y WEBSTER, R. On the pseudo cross-variogram. *Mathematical Geology*, 1993, vol. 25, no 8, p. 1015-1026.
- [17] GOULARD, Michel y VOLTZ, Marc. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 1992, vol. 24, no 3, p. 269-286.
- [18] WACKERNAGEL, Hans. *Multivariate geostatistics: an introduction with applications*. 3rd edition. Springer-Verlag Berlin Heidelberg, 2003.
- [19] GOOVAERTS, Pierre. *Geostatistics for natural resources evaluation*. New York: Oxford University Press, 1997.
- [20] CAICEDO, Eduardo F. y LÓPEZ, Jesús A. *Una aproximación práctica a las redes neuronales artificiales*. Santiago de Cali: Programa Editorial Universidad del Valle, 2009.
- [21] HILERA, José R. y MARTÍNEZ, Víctor J. *Redes Neuronales Artificiales: Fundamentos, modelos y aplicaciones*. Madrid: RA-MA, 1995.
- [22] FLÓREZ, Raquel y FERNÁNDEZ, José M. *Las Redes Neuronales Artificiales: Fundamentos teóricos y aplicaciones prácticas*. Oleiros, La coruña: Netbiblo, 2008.
- [23] FREEMAN, James A.; SKAPURA, David M. *Neural networks: algorithms, applications, and programming techniques*. Massachusetts: Addison-Wesley Publishing Co., Inc., 1991.
- [24] BREIMAN, L., et al. *Classification and Regression Trees*. Belmont: Wadsworth, 1984.
- [25] DÍAZ, Juan F. y CORREA, Juan C. Comparación entre árboles de regresión CART y regresión lineal. *Comunicaciones en Estadística*, 2013, vol. 6, no 2, p. 175-195.
- [26] WU, Xindong, et al. Top 10 algorithms in data mining. *Knowledge and information systems*, 2008, vol. 14, no 1, p. 1-37.
- [27] TAN, Pang-Ning; STEINBACH, Michael y KUMAR, Vipin. Classification: basic concepts, decision trees, and model evaluation. *Introduction to data mining*, 2013, vol. 1, p. 145-205.
- [28] VAPNIK, Vladimir N. *The Nature of Statistical Learning Theory*. New York: Springer, 1995.
- [29] VAPNIK, Vladimir N. *Statistical Learning Theory*. New York: John Wiley and Sons, 1998.
- [30] VAPNIK, Vladimir N. An overview of statistical learning theory. *IEEE transactions on neural networks*, 1999, vol. 10, no 5, p. 988-999

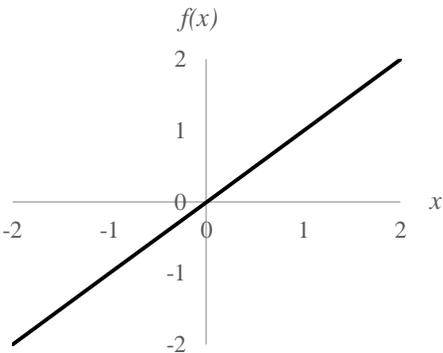
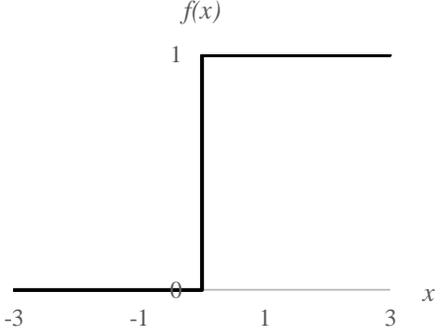
- [31] SMOLA, Alex J.; SCHÖLKOPF, Bernhard. A tutorial on support vector regression. *Statistics and computing*, 2004, vol. 14, no 3, p. 199-222.
- [32] CARMONA SUÁREZ, Enrique J. Tutorial sobre máquinas de Vectores Soporte (SVM). *Madrid: Departamento de Inteligencia Artificial, Universidad Nacional de Educación a distancia (UNED)*, 2016.
- [33] SCHÖLKOPF, Bernhard; SMOLA, Alexander J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press, 2002.
- [34] VAPNIK, Vladimir; GOLOWICH, Steven E.; SMOLA, Alex J. Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. *Advances in Neural Information Processing Systems*, 1997, vol. 9, p. 281-287.
- [35] GUNN, Steve. Support Vector Machines for Classification and Regression. ISIS Technical Report, 1998.
- [36] Joint Ore Reserves Committee (JORC). *Australasian Code for Reporting of Exploration Results, Mineral Resources, and Ore Reserves (The JORC Code)*. 2012.
- [37] ANNELS, Alwyn E. Geostatistical Ore-reserve Estimation. *Mineral Deposit Evaluation*. Springer, Dordrecht, 1991. p. 175-245.
- [38] CORNFORD, Dan. Are comparative studies a waste of time? SIC2004 examined. *Automatic Mapping Algorithms for Routine and Emergency Monitoring Data*, 2005, vol. 1, no 2, p. 61-70.

ANEXO A: SECCIÓN DE REDES NEURONALES ARTIFICIALES

ANEXO A. 1: FUNCIONES DE ACTIVACIÓN PARA LAS ANN

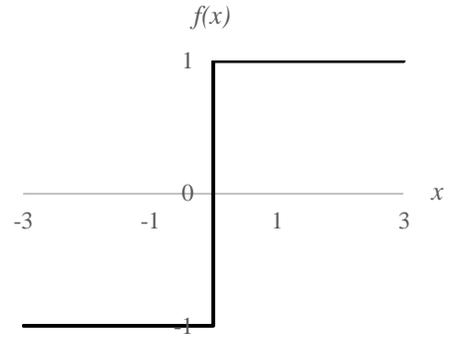
A continuación, se muestra una tabla con la descripción de alguna de las funciones de activación para las redes neuronales [22].

Tabla 47. Principales funciones de activación.

Función	Rango	Gráfica
Identidad	$y = x$	
	$[-\infty, \infty]$	
Escalón	$y = \begin{cases} 1 & \text{si } x \geq c \\ 0 & \text{si } x < c \end{cases}$	$\{0,1\}$
	$c = 0$	

Escalón

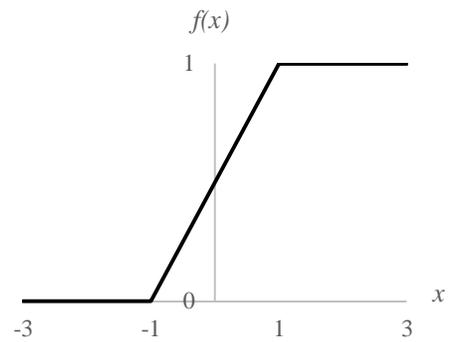
$$y = \begin{cases} 1 & \text{si } x \geq c \\ -1 & \text{si } x < c \end{cases} \quad \{-1,1\}$$



$$c = 0$$

Mixta

$$y = \begin{cases} 1 & \text{si } x \geq C \\ x & \text{si } c \leq x < C \\ 0 & \text{si } x < c \end{cases} \quad [0,1]$$

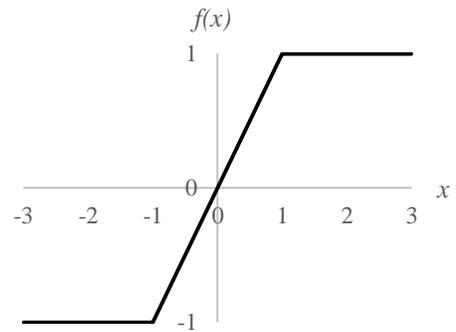


$$c = -1, C = 1$$



Mixta

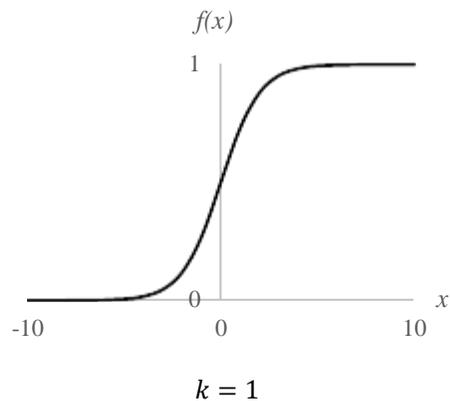
$$y = \begin{cases} 1 & \text{si } x \geq C \\ x & \text{si } c \leq x < C \\ 0 & \text{si } x < c \end{cases} \quad [-1,1]$$



$$c = -1, C = 1$$

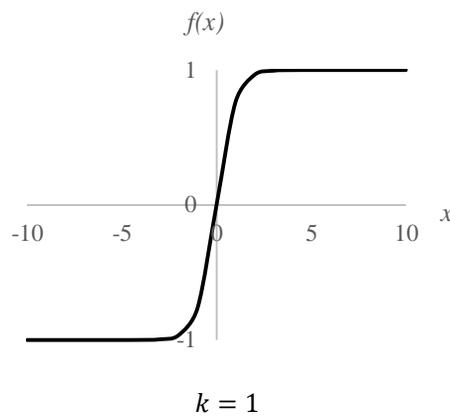
Sigmoidal

$$y = \frac{1}{1 + e^{-kx}}, k > 0 \quad [0,1]$$



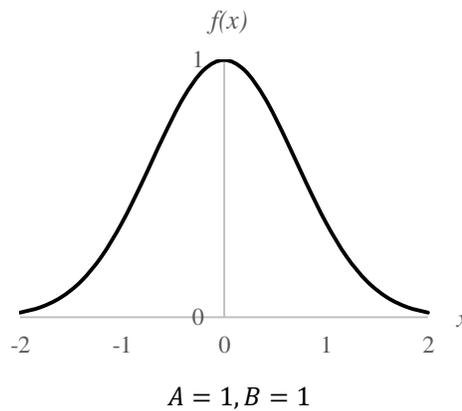
Tangencial

$$y = \frac{e^{kx} - e^{-kx}}{e^{kx} + e^{-kx}}, k > 0 \quad [-1,1]$$



Gaussiana

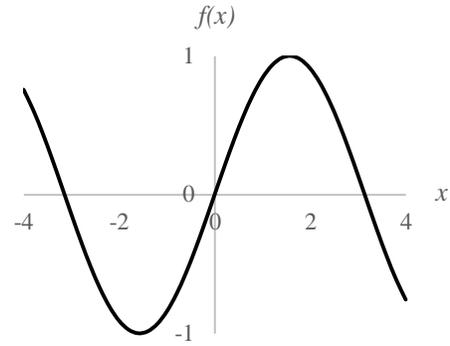
$$y = Ae^{-Bx^2}, A > 0, B > 0 \quad [0,1]$$



Sinusoidal

$$y = A \sin(\omega x + \varphi)$$

$[-1,1]$



$$A = 1, \omega = 1, \varphi = 0$$

ANEXO A. 2: TIPOS DE ANN

En la siguiente tabla se muestra un resumen de los tipos de redes neuronales más conocidos [21], donde se describen sus aplicaciones más importantes, sus ventajas y limitaciones y en qué año y por quien fue creada.

Tabla 48. Tipos de Redes Neuronales Artificiales más importantes.

Modelo de red	Año	Aplicaciones más importantes	Comentarios	Limitaciones	Creador
Avalancha	1967	Reconocimiento de habla continua. Control brazos robot.	Ninguna red sencilla puede hacer todo esto.	No es fácil alterar la velocidad o interpolar el movimiento.	Stephen Grossberg.
Teoría Resonancia Adaptativa (ART)	1986	Reconocimiento de patrones (radar, sonar, etc.).	Sofisticada. Poco utilizada.	Sensible a la translación, distorsión y escala.	Gail Carpenter, Stephen Grossberg.
ADALINE/ MADALINE	1960	Filtrado de señales. Ecuador adaptativo. Modems.	Rápida, fácil de implementar con circuitos analógicos o VLSI.	Sólo es posible clasificar espacios linealmente separados.	Bernard Widrow.
Backpropagation	1974 – 1985	Síntesis de voz desde texto. Control de robots. Predicción.	Red más popular. Numerosas aplicaciones con éxito. Facilidad de	Necesita mucho tiempo para el aprendizaje y muchos ejemplos.	Paul Werbos, David Parker,

		Reconocimiento de patrones.	aprendizaje. Potente.		David Rumelhart.
Memoria Asociativa Bidireccional	1985	Memoria heteroasociativa de acceso por contenido.	Aprendizaje y arquitectura simples.	Baja capacidad de almacenamiento. Los datos deben ser codificados.	Bart Kosko.
Máquinas de Boltzmann y Cauchy	1985 – 1986	Reconocimiento de patrones (imágenes, sonar y radar). Optimización.	Redes simples. Capacidad de representación óptima de patrones.	Necesita un tiempo muy largo de aprendizaje.	Jeffrey Hinton, Terry Sejnowski, Harold Szu.
Brain-Estate-in-a-Box	1977	Extracción de conocimiento de bases de datos.	Posiblemente mejor realización que las redes de Hopfield.	Realización y potenciales aplicaciones no estudiadas totalmente.	James Anderson.
Cerebellatron	1969	Control del movimiento de los brazos de un robot.	Semejante a Avalancha.	Requiere complicadas entradas de control.	David Marr, James Albus, Andres Pellionez.
Counter-propagation	1986	Comprensión de imágenes.	Combinación de Perceptron y TPM.	Numerosas neuronas y conexiones.	Robert Hecht-Nielsen.
Hopfield	1982	Reconstrucción de patrones y optimización.	Puede implementarse en VLSI. Fácil de conceptualizar.	Capacidad y estabilidad.	John Hopfield.
Neocognitron	1978 – 1984	Reconocimiento de caracteres manuscritos.	Insensible a la translación, rotación y escala.	Requiere muchos elementos de proceso, niveles y conexiones.	K. Fukushima.
Perceptron	1957	Reconocimiento de caracteres impresos.	La red más antigua. Construida en HW.	No puede reconocer caracteres complejos.	Frank Rosenblatt.
Self-Organizing-Map (SOM), Topology-Preserving-Map (TPM).	1980 – 1984	Reconocimiento de patrones, codificación de datos, optimización.	Realiza mapas de características comunes de los datos aprendidos.	Requiere mucho entrenamiento.	Teuvo Kohonen.

ANEXO B: SECCIÓN DE SVM

ANEXO B. 1: INTRODUCCIÓN A LA TEORÍA DE OPTIMIZACIÓN

Sea el siguiente problema de optimización denominado problema primal:

$$\begin{aligned} &\text{Minimizar} && f(x), \quad x \in \Omega \\ &\text{Sujeto a} && g_i(x) \leq 0, \quad \forall i = 1, \dots, n \end{aligned} \tag{87}$$

La solución del problema anterior, x^* , cumple que $g_i(x^*) \leq 0$ y $f(x^*) \leq f(x)$, $\forall x$ tal que $g_i(x) \leq 0$. Si la función a optimizar y sus restricciones son lineales, entonces el problema se aborda mediante programación lineal. Si, por el contrario, la función a optimizar es cuadrática, con las mismas condiciones de restricciones lineales, el problema se aborda utilizando programación cuadrática. La Teoría de Optimización dicta que un problema primal tiene su forma dual siempre que la función a optimizar y las restricciones sean estrictamente convexas. Para realizar esta conversión, se utiliza la función de Lagrange, definida como:

$$L(x, \alpha) = f(x) + \sum_{i=1}^n \alpha_i g_i(x), \tag{88}$$

donde $\alpha_i \geq 0$ son los multiplicadores de Lagrange. A partir de la función anterior, se define el problema dual:

$$\begin{aligned} &\text{Maximizar} && \varphi(\alpha) = \inf_{x \in \Omega} L(x, \alpha) \\ &\text{Sujeto a} && \alpha_i \geq 0, \quad \forall i = 1, \dots, n \end{aligned} \tag{89}$$

El interés de hacer dicha transformación es que la resolución del problema dual es más fácil que el problema primal. Además, al resolver el problema dual, bajo ciertas condiciones, también se resuelve el problema primal.

ANEXO B. 2: TRANSFORMACIÓN A PROBLEMA DUAL

- **Paso 1:** Obtención de la función lagrangiana.

$$L = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*) - \sum_{i=1}^n \alpha_i (\varepsilon + \xi_i - y_i + \langle \mathbf{w}, \mathbf{x}_i \rangle + b) - \sum_{i=1}^n \alpha_i^* (\varepsilon + \xi_i^* + y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b), \quad (90)$$

donde $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ son multiplicadores de Lagrange. Además, las variables duales tienen que satisfacer la condición de ser positivas, es decir:

$$\eta_i, \eta_i^*, \alpha_i, \alpha_i^* \geq 0, \quad \forall i = 1, \dots, n \quad (91)$$

- **Paso 2:** Aplicación de las condiciones de KKT.

$$\frac{\partial L}{\partial b} = 0 \quad (92)$$

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \quad (93)$$

$$\frac{\partial L}{\partial \xi_i} = 0 \quad (94)$$

$$\frac{\partial L}{\partial \xi_i^*} = 0 \quad (95)$$

$$\alpha_i (\varepsilon + \xi_i - y_i + \langle \mathbf{w}, \mathbf{x}_i \rangle + b) = 0 \quad (96)$$

$$\alpha_i^* (\varepsilon + \xi_i^* + y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b) = 0 \quad (97)$$

$$\eta_i \xi_i = 0 \quad (98)$$

$$\eta_i^* \xi_i^* = 0 \quad (99)$$

- **Paso 3:** Establecer las relaciones entre las variables del problema primal (w, b, ξ_i, ξ_i^*) con la del problema dual ($\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$), por lo que es necesario utilizar la ecuación (93).

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i \quad (100)$$

- **Paso 4:** Establecer restricciones adicionales de las variables duales. Se hace uso de las ecuaciones (92), (94) y (95).

$$\sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \quad (101)$$

$$\eta_i = C - \alpha_i \quad (102)$$

$$\eta_i^* = C - \alpha_i^* \quad (103)$$

- **Paso 5:** De lo obtenido en el paso 3, se eliminan las variables primales de la función lagrangiana.

$$L = -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \quad (104)$$

ANEXO B. 3. CÁLCULO DE b

De las condiciones de KKT descritas anteriormente, se puede desprender que:

$$(C - \alpha_i) \xi_i = 0 \quad (105)$$

$$(C - \alpha_i^*) \xi_i^* = 0 \quad (106)$$

Suponiendo que el ejemplo (x_i, y_i) se encuentra dentro de la zona tubular ε -insensible, implicando que $\xi_i = 0$ y $\xi_i^* = 0$, y no pertenece a la frontera, entonces los factores de las ecuaciones (96) y (97) son distinto de cero, por lo que se debe cumplir que $\alpha_i = 0$ y $\alpha_i^* = 0$, respectivamente.

Por otro lado, si $\alpha_i > 0$, el factor de la ecuación (96) debe ser nulo, es decir:

$$\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b = 0 \quad (107)$$

Además, si $\alpha_i < C$ entonces de la ecuación (105) se desprende que:

$$\xi_i = 0 \quad (108)$$

Combinando los resultados anteriores se obtiene lo siguiente:

$$\alpha_i(\varepsilon - y_i + \langle \mathbf{w}, \mathbf{x}_i \rangle + b) = 0, \quad \text{si } 0 < \alpha_i < C \quad (109)$$

De la misma forma, si $\alpha_i^* > 0$, el factor de la ecuación (97) es nulo:

$$\varepsilon + \xi_i^* + y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b = 0 \quad (110)$$

Si $\alpha_i^* < C$, de la ecuación (106) se infiere que:



$$\xi_i^* = 0 \quad (111)$$

Al combinar las ecuaciones (110) y (111) se obtiene:

$$\alpha_i^*(\varepsilon + y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - b) = 0, \quad \text{si } 0 < \alpha_i^* < C \quad (112)$$

ANEXO C: LEY EQUIVALENTE Y LEY DE CORTE

ANEXO C. 1: LEY EQUIVALENTE

La ley equivalente se utiliza cuando coexisten diversos tipos de minerales en un mismo yacimiento, caracterizando a la ley del mineral de interés como una combinación aditiva entre ella y los subproductos:

$$Ley_{eq} = Ley_{int} + \sum_{i=1}^N F_{eq,i} \times Ley_{sub,i}, \quad (113)$$

donde los subíndices *int* y *sub* se refieren al mineral de interés y a los subproductos, respectivamente, y $F_{eq,i}$ se refiere al factor de conversión del *i*-ésimo subproducto. Los ponderadores dependen de factores económicos y de la recuperación, expresado en la siguiente ecuación:

$$F_{eq} = \frac{(s_{sub} - r_{sub}) \times y_{sub}}{(s_{int} - r_{int}) \times y_{int}}, \quad (114)$$

donde *s* es el precio de venta del mineral, definida por la ecuación (115), *r* es el costo de venta, que incluye el costo de tratamiento, refinación y transporte expresado en la ecuación (116), e *y* es la recuperación del mineral.

$$s = Precio \times \%pagable \quad (115)$$

$$r = Costo\ tratamiento + Costo\ refinacion + Costo\ transporte \quad (116)$$

ANEXO C. 2: LEY DE CORTE

La ley de corte se define como la ley más baja económicamente extraíble de un determinado yacimiento, donde toda ley sobre esta se define como mineral y es enviado a planta para ser procesado, mientras que las leyes que se encuentren bajo esta son definidas como estéril y se envían a botaderos.

El cálculo de la ley de corte en términos prácticos se define según la ecuación (117).

$$COG = \frac{CM - CP}{(s - r) \times y'} \quad (117)$$

donde CM es el costo mina y CP es el costo de procesamiento del mineral de interés.



ANEXO D: CASO SIMULADO

A continuación, se muestran los resultados obtenidos de error de estimación y beneficio económico para los cien casos simulados.

ANEXO D. 1: SKILL VALUE

En la siguiente tabla se especifican los valores obtenidos para el *skill value* para los cuatro métodos de estimación utilizados en las cien simulaciones.

Tabla 49. *Skill values* para cobre, molibdeno y oro en las cien simulaciones.

N°	Cobre				Molibdeno				Oro			
	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR	OCOK	ANN	CART	SVR
1	0.2493	0.2603	0.2854	0.2392	24.13	24.54	26.85	22.44	0.2408	0.2508	0.2472	0.2373
2	0.2349	0.2398	0.2557	0.2208	22.74	23.19	25.44	21.92	0.246	0.262	0.252	0.2288
3	0.2424	0.259	0.2661	0.2493	23.98	24.75	25.02	24.55	0.2567	0.2466	0.2809	0.2441
4	0.2687	0.2546	0.3043	0.2526	26.32	25.06	31.28	24.22	0.252	0.234	0.2493	0.224
5	0.237	0.2271	0.2594	0.2233	26.41	24.69	29.74	25.03	0.2496	0.2445	0.2525	0.2354
6	0.2413	0.2406	0.2777	0.2292	26.06	25.55	28.4	24.14	0.2632	0.2605	0.2689	0.239
7	0.2845	0.2725	0.2838	0.2531	28.88	27.62	32.05	24.72	0.2606	0.2497	0.2806	0.2341
8	0.2977	0.2674	0.2883	0.2642	27.37	25.9	29.17	24.04	0.2604	0.2584	0.2898	0.2372
9	0.2508	0.2462	0.2529	0.2457	24.18	26.25	26.44	24.34	0.2344	0.2452	0.2873	0.226
10	0.2595	0.2379	0.2759	0.245	26.16	25.19	28.75	24.68	0.2407	0.2322	0.2423	0.2362
11	0.2436	0.2333	0.2472	0.2235	29.86	27.51	30.83	26.9	0.2904	0.2734	0.2956	0.2614
12	0.2888	0.2588	0.3041	0.2422	26.95	23.7	28.33	24.51	0.2657	0.2473	0.2674	0.2497
13	0.2537	0.2528	0.2798	0.242	24.43	24.54	27.57	22.89	0.2593	0.2644	0.2695	0.2548
14	0.2521	0.2568	0.2736	0.2296	22.9	22.19	23	21.14	0.247	0.2505	0.2688	0.2262
15	0.2435	0.2392	0.2507	0.2251	29.77	28.82	31.52	26.67	0.243	0.2394	0.2597	0.2208
16	0.2426	0.2361	0.2471	0.2262	25.06	24.26	26.71	23.24	0.2507	0.2377	0.268	0.2344
17	0.2765	0.2619	0.2827	0.2453	28.09	27.18	30.55	24.46	0.2661	0.2525	0.2824	0.2466
18	0.2412	0.2321	0.2677	0.2236	26.51	24.63	28.25	23.31	0.2675	0.2514	0.287	0.2583
19	0.2751	0.2488	0.2832	0.2422	27.69	27.07	28.97	25.46	0.2739	0.2481	0.2915	0.2538
20	0.2516	0.2619	0.264	0.2234	26.26	26.05	28.46	24.33	0.2547	0.2402	0.2494	0.2368
21	0.2566	0.2502	0.2824	0.24	23.04	21.75	24.83	21.1	0.2641	0.2427	0.265	0.2318
22	0.2579	0.243	0.2764	0.2384	28.14	26.22	28.4	26.04	0.2882	0.2903	0.2969	0.2556
23	0.2702	0.265	0.3109	0.2506	23.13	22.74	23.51	22.6	0.2467	0.2401	0.2731	0.2327
24	0.2678	0.244	0.2943	0.2376	25.29	25.66	25.59	23.14	0.2498	0.2454	0.2859	0.229

25	0.2846	0.2514	0.3088	0.2488	25.55	24.12	27.6	24.46	0.2409	0.2178	0.2437	0.2204
26	0.2671	0.2478	0.2876	0.244	25.42	24.27	27.24	23.44	0.2419	0.2407	0.2659	0.2382
27	0.2405	0.2468	0.2761	0.228	25.7	23.29	29.53	24.15	0.237	0.2486	0.237	0.2251
28	0.2639	0.2557	0.277	0.2485	26.03	24.73	26.16	23.61	0.2659	0.256	0.2776	0.2398
29	0.238	0.2409	0.246	0.224	23.51	25.06	26.64	22.43	0.2339	0.2424	0.2636	0.2167
30	0.2459	0.2423	0.2649	0.2381	24.15	23.91	25.26	22.52	0.2519	0.2513	0.2804	0.2365
31	0.3253	0.3083	0.3169	0.2544	31.24	28.99	31.1	24.46	0.3141	0.2952	0.3278	0.2467
32	0.2606	0.2455	0.2742	0.2357	21.87	21.73	21.7	21.07	0.2318	0.2487	0.259	0.2373
33	0.2509	0.2755	0.2546	0.2383	26.48	28.78	26.29	23.93	0.2253	0.2533	0.2238	0.1996
34	0.2697	0.2679	0.2787	0.2484	25.68	24.21	28.12	23.82	0.2557	0.2421	0.2569	0.2368
35	0.2484	0.2441	0.2448	0.2302	24.54	24.41	26.12	22.94	0.2659	0.2671	0.2841	0.2574
36	0.2763	0.2552	0.2762	0.2495	25.89	25.94	25.88	23.63	0.2772	0.2807	0.293	0.2515
37	0.2587	0.2534	0.2692	0.239	27.86	28.88	29.44	24.86	0.25	0.2618	0.2682	0.2357
38	0.2317	0.2316	0.2413	0.223	26.59	25.26	28.19	24.95	0.2526	0.2465	0.3085	0.2292
39	0.2487	0.2448	0.2624	0.2364	25.08	24.59	26.82	23.18	0.245	0.2465	0.2748	0.2425
40	0.2291	0.2553	0.2556	0.2235	21.79	21.98	21.8	21.06	0.2353	0.2456	0.2604	0.2232
41	0.2405	0.2281	0.2457	0.2485	24.63	24.68	25.47	24.14	0.2454	0.2387	0.2724	0.2403
42	0.2448	0.264	0.2594	0.2236	26.23	26.29	26.1	22.22	0.2605	0.2536	0.2462	0.2307
43	0.2611	0.2687	0.2616	0.247	26.87	26.87	28.51	25.77	0.2289	0.2424	0.2437	0.2255
44	0.2652	0.2727	0.2738	0.2489	23.68	24.63	24.99	23.18	0.2658	0.2565	0.2716	0.2518
45	0.279	0.2648	0.3114	0.2505	27.11	25.51	29.25	24.21	0.262	0.2532	0.2648	0.2427
46	0.2734	0.2476	0.2819	0.2483	24.28	23.32	23.91	22.88	0.2676	0.257	0.2857	0.2526
47	0.2539	0.2617	0.274	0.2429	24.2	23.69	24.58	23.11	0.238	0.2335	0.259	0.2238
48	0.262	0.2573	0.2688	0.2316	28	26.27	26.99	23.7	0.2749	0.2658	0.2858	0.2396
49	0.2549	0.2569	0.2593	0.2316	23.5	25.18	25.13	22.61	0.2703	0.2523	0.2896	0.2445
50	0.266	0.2538	0.267	0.2504	25.1	26.34	27.13	24.87	0.2819	0.2487	0.2597	0.2396
51	0.2516	0.2474	0.2956	0.2422	23.67	22.84	26.08	23.22	0.264	0.2602	0.2676	0.2557
52	0.2732	0.2627	0.2825	0.2535	24.52	23.69	24.97	22.94	0.2698	0.2569	0.2967	0.2454
53	0.2568	0.2483	0.2798	0.2479	25.53	24.94	28.4	24.73	0.2735	0.27	0.268	0.2566
54	0.2586	0.264	0.27	0.2389	23.46	22.71	23.87	22.58	0.2381	0.232	0.2541	0.2359
55	0.2403	0.242	0.2672	0.228	25.13	24.13	25.12	23.54	0.2607	0.2527	0.2663	0.2446
56	0.2559	0.2609	0.2634	0.2475	26.65	26.63	27.86	25.92	0.2461	0.2522	0.2589	0.2305
57	0.2816	0.2684	0.2713	0.2748	27.85	24.9	30.26	26.37	0.2817	0.2449	0.2856	0.252
58	0.2727	0.2801	0.2735	0.247	24.49	27.17	25.67	23.87	0.2507	0.2575	0.2737	0.2467
59	0.2615	0.2517	0.266	0.234	24.63	24.54	25.09	24.16	0.2491	0.2377	0.2658	0.2332
60	0.2517	0.2315	0.2777	0.2246	28.9	25.28	27.66	24.46	0.2861	0.2797	0.3093	0.2577
61	0.2279	0.2203	0.2195	0.2172	23.08	22.32	24.64	22.01	0.2331	0.2271	0.2532	0.2291
62	0.2384	0.2394	0.258	0.2262	22.81	22.07	25.78	21.65	0.2152	0.2143	0.2164	0.2032
63	0.2376	0.2418	0.248	0.2313	22.66	23.55	23.16	21.97	0.2242	0.2389	0.2526	0.2203
64	0.2349	0.2374	0.2459	0.2279	23.1	26.58	26.62	23.06	0.2325	0.2557	0.2591	0.2295
65	0.2486	0.2415	0.2429	0.2282	28.5	24.84	30.13	25	0.2589	0.2753	0.2704	0.2531
66	0.2357	0.2358	0.2682	0.2236	25.42	25.38	26.4	24.99	0.2427	0.2361	0.2759	0.2318

67	0.2619	0.2601	0.2912	0.2399	26.09	25.25	26.54	24	0.2612	0.2589	0.3106	0.2635
68	0.2466	0.256	0.2926	0.2359	25.06	25.92	25.56	23.23	0.239	0.2499	0.2892	0.2346
69	0.2287	0.2434	0.2572	0.2275	25.53	25.38	25.95	24.85	0.2442	0.2551	0.2955	0.2358
70	0.2629	0.2614	0.2917	0.2426	25.7	27	27.38	24.77	0.2624	0.261	0.2624	0.2446
71	0.2512	0.2409	0.2651	0.2453	25.8	24.41	26.79	24.55	0.245	0.2454	0.2424	0.2356
72	0.239	0.2595	0.2584	0.2385	24.38	24.83	25.09	24.15	0.2454	0.2622	0.276	0.2353
73	0.2446	0.2409	0.2557	0.2409	26.59	26.01	29.69	25.76	0.2487	0.2498	0.2828	0.2497
74	0.2302	0.2327	0.2341	0.2318	26.64	26.64	29.3	25.34	0.2562	0.258	0.2862	0.245
75	0.2335	0.2377	0.236	0.2337	21.7	23.54	23.46	21.94	0.2373	0.242	0.267	0.2318
76	0.2674	0.261	0.273	0.2479	25.61	25.33	26.97	25.22	0.2389	0.2458	0.2605	0.2349
77	0.2594	0.262	0.2681	0.2319	26.67	27.14	27.47	24.43	0.2725	0.2596	0.275	0.2414
78	0.2624	0.251	0.2846	0.2463	27.65	26.56	27.5	25.75	0.2655	0.2824	0.2778	0.2648
79	0.2726	0.2503	0.267	0.2608	33.13	27.63	32.29	29.45	0.2759	0.2543	0.2948	0.2705
80	0.2694	0.2657	0.2652	0.2462	28.53	25.19	28.14	24.88	0.2613	0.2499	0.2645	0.2411
81	0.2369	0.2334	0.2409	0.217	26.72	25.25	26.45	23.4	0.2564	0.2439	0.2585	0.2313
82	0.2244	0.2225	0.2516	0.2144	21.13	20.87	21.65	20.12	0.1991	0.2049	0.2217	0.1927
83	0.2476	0.2585	0.3055	0.2397	25.21	23.68	28.83	22.75	0.2248	0.2205	0.2279	0.2155
84	0.2502	0.2445	0.2647	0.2379	27.21	26.11	28.36	25.4	0.2698	0.2695	0.2683	0.2567
85	0.2857	0.2685	0.306	0.2632	29.84	28.26	30.2	26.53	0.2705	0.2448	0.2773	0.2353
86	0.252	0.2558	0.2728	0.2386	27.05	27.64	27.66	26.08	0.2501	0.2518	0.2756	0.24
87	0.2412	0.2492	0.2655	0.2392	24.82	27.01	25.32	26.16	0.2692	0.2848	0.3049	0.268
88	0.283	0.245	0.3056	0.2375	28.16	23.65	30.09	23.24	0.2792	0.2554	0.2777	0.266
89	0.2803	0.2512	0.2758	0.2434	30.2	28	30.63	24.65	0.2813	0.2669	0.3064	0.2433
90	0.2346	0.2411	0.2803	0.2249	23.25	22.61	27.09	22.11	0.2451	0.2353	0.2467	0.2288
91	0.2239	0.2365	0.2457	0.2181	22.92	23.47	23.47	21.97	0.2423	0.2397	0.2607	0.2246
92	0.2466	0.2361	0.2535	0.2209	26.13	24.27	26.15	23.11	0.2414	0.217	0.2504	0.2098
93	0.2665	0.2574	0.2632	0.2468	28.3	30.43	30.43	26.23	0.2557	0.2537	0.2762	0.2409
94	0.2668	0.2864	0.2846	0.2721	24.9	25.79	25.75	23.88	0.249	0.2607	0.2834	0.2329
95	0.2806	0.2622	0.3051	0.2449	25.72	24.97	27.32	23.48	0.2864	0.2765	0.2812	0.2354
96	0.2343	0.2404	0.2868	0.2214	25.32	25.67	27.11	24.45	0.2376	0.2282	0.2483	0.227
97	0.2535	0.2636	0.2878	0.2559	23.22	22.82	25.96	22.04	0.2776	0.2753	0.2902	0.2475
98	0.2583	0.2694	0.2873	0.2495	24.86	24.76	25.72	24.34	0.2483	0.251	0.2723	0.2292
99	0.229	0.2373	0.225	0.2176	26.26	22.15	26.2	23.63	0.2608	0.2468	0.2854	0.2447
100	0.2589	0.2434	0.2664	0.2637	25.08	24.66	25.58	23.51	0.2612	0.2534	0.2621	0.2563

ANEXO D. 2: EVALUACIÓN ECONÓMICA

En la Tabla 50 se encuentran las diferencias entre el beneficio económico real obtenido de la base original y el obtenido con la estimación realizada para las cien simulaciones.

Tabla 50. Diferencia entre los beneficios de la base original y la base estimada en las cien simulaciones.

N°	OCOK [MUSD\$]	ANN [MUSD\$]	CART [MUSD\$]	SVR [MUSD\$]	N°	OCOK [MUSD\$]	ANN [MUSD\$]	CART [MUSD\$]	SVR [MUSD\$]
1	-2.74	1.48	-6.27	0.83	33	-7.41	-15.50	-5.01	1.25
2	-7.40	-11.54	-9.14	-2.36	34	-2.30	-8.73	0.28	-1.28
3	2.43	-4.69	2.28	3.16	35	-6.46	-1.87	-6.12	-3.01
4	-6.10	-4.96	-2.56	-1.00	36	-15.67	3.60	-11.64	-4.21
5	6.14	-2.36	7.47	0.96	37	-9.44	-10.64	-5.21	-0.82
6	-6.61	-8.21	-15.34	-2.10	38	7.74	1.94	14.97	0.86
7	10.56	2.11	12.45	1.22	39	1.09	3.83	7.89	2.07
8	-11.94	-8.25	-4.86	-0.82	40	-0.80	13.25	-5.69	3.05
9	-5.96	1.88	-14.98	-5.89	41	-2.15	-2.76	6.63	-0.90
10	5.82	-0.27	10.62	5.40	42	-8.32	-11.09	-6.17	0.49
11	9.81	-1.87	10.70	1.19	43	2.14	6.27	2.29	4.20
12	9.35	-7.12	12.63	7.61	44	10.30	9.60	8.50	5.38
13	5.23	1.13	8.39	0.46	45	10.44	-1.43	16.82	4.65
14	-12.01	-14.73	-13.46	-2.97	46	12.80	-4.57	10.33	1.75
15	6.01	6.87	8.83	3.67	47	4.13	1.17	-0.27	-0.93
16	0.83	-1.69	2.23	1.20	48	8.66	7.61	13.02	0.69
17	-14.65	-8.52	-17.11	-6.46	49	9.82	2.07	10.99	4.18
18	-2.88	-1.25	-8.16	-4.90	50	-3.31	3.35	5.71	-0.48
19	-13.26	-1.78	-10.84	-4.30	51	3.80	3.09	10.18	-1.25
20	12.81	6.92	9.92	3.52	52	-9.20	0.36	-10.30	-1.07
21	6.62	-2.78	8.18	4.32	53	-4.17	-1.05	-3.90	-2.28
22	9.51	7.69	9.59	1.89	54	7.05	3.27	8.70	6.93
23	-6.97	2.78	-10.12	0.74	55	0.98	5.05	4.10	0.16
24	-12.30	3.74	-17.25	-1.28	56	-3.85	-5.85	3.26	-2.27
25	10.01	1.49	13.01	6.30	57	13.83	6.27	15.01	9.91
26	6.32	-0.51	12.33	5.27	58	-9.17	-9.13	-11.17	-3.54
27	-1.84	-11.40	4.60	1.49	59	2.47	0.20	-1.45	2.08
28	-4.95	-9.11	-7.45	4.63	60	-13.42	-10.55	-19.13	-3.91
29	-4.19	5.26	-7.57	0.37	61	-3.39	-1.67	-5.41	2.85
30	7.34	7.22	14.20	5.14	62	2.97	1.22	3.48	-0.15
31	26.06	18.06	24.82	6.99	63	2.56	4.22	-3.24	0.80
32	9.99	-10.87	11.14	3.50	64	-1.08	-10.27	-7.13	-0.70

65	5.71	-4.89	7.76	1.58	83	-0.84	9.61	-8.09	4.71
66	-2.39	-0.69	-1.28	4.02	84	-6.60	-7.13	-6.74	-0.43
67	9.43	3.96	17.83	4.56	85	-19.39	-4.73	-15.42	-8.47
68	4.35	-1.37	14.23	4.60	86	4.66	7.90	4.36	3.32
69	1.92	2.64	3.84	1.40	87	-3.29	2.92	-6.50	-4.99
70	5.11	-8.42	5.72	-1.95	88	-14.45	5.40	-12.79	-3.77
71	6.35	-2.97	1.53	3.37	89	12.76	12.09	15.81	4.48
72	3.22	4.52	7.39	2.93	90	0.88	-4.82	7.44	2.82
73	-5.19	-7.37	-10.94	1.99	91	4.69	-6.88	7.51	3.39
74	-5.77	3.88	-9.30	2.44	92	-9.10	2.85	-4.08	-0.14
75	-0.35	0.95	-0.05	0.82	93	-1.22	-0.18	-2.27	6.48
76	3.39	5.24	7.53	1.12	94	6.11	12.67	8.66	5.16
77	-12.46	-8.72	-8.83	-2.61	95	13.10	10.81	14.41	3.10
78	-5.16	-1.29	-1.45	2.98	96	-7.01	-5.54	-14.34	-1.29
79	-11.17	-5.27	-11.96	-1.85	97	9.88	2.52	13.01	8.33
80	-11.66	2.73	-5.14	-0.99	98	-1.05	-0.89	-7.69	1.94
81	-3.25	-0.32	3.32	1.67	99	4.12	7.59	9.23	3.37
82	-0.07	-4.02	-4.13	0.28	100	0.60	0.60	6.60	-3.13



ANEXO E: CASO REAL

En esta sección se muestran los resultados obtenidos del caso real según cada dominio.

ANEXO E. 1: EVALUACIÓN ECONÓMICA

La Tabla 51 contiene el ingreso, costo y beneficio obtenido con cada método de estimación empleado en el trabajo de investigación en cada dominio.

Tabla 51. Evaluación económica para estimación de bloques según dominio.

		Ingreso [MUSD\$]			Costo [MUSD\$]			Beneficio [MUSD\$]		
		Total	Medidos	Indicados	Total	Medidos	Indicados	Total	Medidos	Indicados
Dominio 1	OCOK	1,329	51.63	1,277	821.7	23.91	797.8	507.1	27.72	479.3
	ANN	1,046	34.83	1,011	591.4	19.84	571.6	454.5	14.99	439.5
	CART	1,241	52.24	1,189	827.1	23.91	803.2	414.0	28.33	385.7
	SVR	1,087	44.28	1,042	838.8	23.91	814.9	247.7	20.37	227.4
Dominio 2	OCOK	1,793	74.50	1,719	1094.9	31.88	1,063	698.2	42.62	655.6
	ANN	1,667	25.75	1,642	831.9	16.79	815.1	835.4	8.97	826.4
	CART	1,900	72.21	1,828	1,112	31.88	1,081	787.5	40.33	747.2
	SVR	1,370	55.13	1,315	1,116	31.71	1,085	253.4	23.43	230.0
Dominio 3	OCOK	4,228	125.5	4,102	1,297	26.45	1,271	2,930	99.07	2,831
	ANN	3,703	120.1	3,583	1,246	26.45	1,219	2,457	93.62	2,363
	CART	4,062	115.0	3,947	1,298	26.45	1,272	2,764	88.57	2,676
	SVR	4,455	122.0	4,333	1,299	26.45	1,272	3,157	95.52	3,061
Dominio 4	OCOK	2,662	21.42	2,641	1,525	8.48	1,517	1,137	12.94	1,124
	ANN	2,491	17.43	2,474	1,334	8.48	1,326	1,157	8.95	1,148
	CART	2,537	20.30	2,517	1,477	8.48	1,468	1,061	11.82	1,049
	SVR	2,661	18.48	2,643	1,585	8.48	1,577	1,076	10.00	1,066

UNIVERSIDAD DE CONCEPCIÓN - FACULTAD DE INGENIERÍA
Departamento de Ingeniería Metalúrgica
 Hoja Resumen Memoria de Título

Título: Estimación polimetálica empleando técnicas de Inteligencia Computacional	
Nombre Memorista: Pamela Antonia Mella Figueroa	
Modalidad	Profesor (es) Patrocinante (s)
Concepto	Prof. Roberto Fustos T.
Calificación	
Fecha	
Prof. FERNANDO PARADA	
Ingeniero Supervisor:	
Institución:	
Comisión (Nombre y Firma)	
Prof. Asieh Hekmat	Prof. Jean Paul Navarrete
Resumen	
<p>La estimación de recursos minerales es esencial para las etapas posteriores de un proyecto minero, pues una mala ejecución de esta conlleva a pérdidas monetarias significativas. Técnicas geoestadísticas tradicionales de estimación son las populares para ser usadas en la industria minera, sin embargo, cuando el yacimiento presenta un comportamiento geológico complejo, los resultados obtenidos en esta no son los mejores. Los recientes avances que han presentado las técnicas de Inteligencia Computacional han hecho de este un foco de atención, provocando una serie de investigaciones para la estimación de leyes minerales. Se propone utilizar los algoritmos Redes Neuronales Artificiales, Árboles de Regresión y <i>Support Vector Regression</i> para ser comparados con el método tradicional Cokriging Ordinario en situaciones donde se tenga baja correlación espacial, poca información de los datos y una alta correlación entre variables.</p> <p>La metodología propuesta consta en dos casos de estudio: uno simulado y uno real. El primero consiste en realizar cien simulaciones gaussianas de tres variables continuas pertenecientes a un yacimiento, donde cada base generada es particionada en una base para entrenar los algoritmos y otra para realizar la validación de estos. El segundo caso corresponde a un yacimiento real, que contiene dos variables continuas, oro y cobre, además de cuatro dominios geometalúrgicos. Una vez obtenido los resultados en los dos casos, se realiza la comparación de los métodos en cada caso individualmente y luego entre uno y otro. Para ambos casos se realiza un plan de categorización y se mide la calidad de esta, finalizando con una evaluación económica.</p> <p>Los resultados obtenidos indican que el método <i>Support Vector Regression</i> obtiene mejores resultados en cuanto a la disminución del error de estimación en comparación con el método tradicional. Para los casos simulados, el 95% de las veces <i>Support Vector Regression</i> obtiene mejores resultados que Cokriging Ordinario, disminuyendo el error de estimación en aproximadamente un 7%, mientras que, para el caso real, en el dominio en que se replican las condiciones señaladas en un inicio, se logra una disminución del 8%. Además, este método junto a Árboles de Regresión, parecen ser los más conservadores a la hora de estimar.</p>	