UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA CIVIL

MODELACIÓN DE LA VARIABILIDAD ESPACIAL DE LA PROFUNDIDAD DE NIEVE CON MACHINE LEARNING Y GEOESTADÍSTICA UTILIZANDO DATOS LIDAR EN CUENCA PIUQUENES, REGIÓN METROPOLITANA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL

ÁLVARO ANTONIO JORQUERA TAPIA

PROFESOR GUÍA:
JAMES MCPHEE TORRES

MIEMBROS DE LA COMISIÓN:
XAVIER EMERY
PABLO MENDOZA ZÚÑIGA

SANTIAGO DE CHILE
AÑO 2019

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL
POR: ÁLVARO ANTONIO JORQUERA TAPIA
FECHA: AÑO 2019
PROF. GUÍA: JAMES MCPHEE TORRES

MODELACIÓN DE LA VARIABILIDAD ESPACIAL DE LA PROFUNDIDAD DE NIEVE CON MACHINE LEARNING Y GEOESTADÍSTICA UTILIZANDO DATOS LIDAR EN CUENCA PIUQUENES, REGIÓN METROPOLITANA

En el presente trabajo de título se desarrollan modelos que caractericen la variabilidad espacial de la profundidad de nieve mediante modelos de Machine Learning (ML) y Geoestadística. Se utilizan 3 fechas de información LiDAR terrestre obtenidas en campañas de terreno en la cuenca experimental Piuquenes durante la temporada 2018.

Para ambos tipos de modelos, se utilizan variables topográficas como variables independientes. Las variables utilizadas son la altura, pendiente, curvatura del terreno, Upwind Slope (Sx) y el Topographic Position Index (TPI). Las variables independientes se obtienen a partir de un DEM tomado con LiDAR durante el verano (resolución de 5 metros). En los modelos, TPI muestra ser la variable independiente más influyente para 2 de las 3 fechas LiDAR.

Se crean tres modelos de ML: Redes Neuronales (ANN), Random Forest (RF) y Gradient-Boost (GBR). Los modelos de RF y GBR están basados en árboles de regresión, método ampliamente usado para explicar la variabilidad de la profundidad de nieve. De los tres modelos, Random Forest muestra el mejor comportamiento pudiendo explicar un 44-48 % de la variabilidad de la profundidad de nieve. Del análisis de los modelos de ML se llega a la conclusión que el parámetro menos influyente en los modelos es el Sx, lo que se explica debido a que no se tiene información de la profundidad de nieve en sitios donde Sx es negativo, razón por la cual la variable no logra tomar importancia en los modelos.

Se modela además la profundidad de nieve mediante el método de co-kriging utilizando las mismas variables independientes antes mencionadas. Se modela el co-kriging utilizando validación cruzada con el método *leave one out*. El análisis de los estadísticos de los errores de los modelos muestra que, en general, co-kriging entrega un desempeño parecido a los métodos de ML, exceptuando una fecha donde co-kriging es capaz de explicar el 76 % de la varianza de la variable de interés.

Agradecimientos

Quiero agradecer en primera parte con nombre y apellido a mi padre Mario Jorquera y mi hermano Sebastian Jorquera por las discusiones en el auto, en la casa, por interesarse en mi trabajo, preguntar, opinar y ayudar. Por hacer de mi tesis un tema de conversación en el cual se me planteaban preguntas a diario, muchas veces difíciles de contestar, muchas gracias, sin su aliento hubiera sido todo mucho mas difícil. A mi familia y amigos de la universidad, del colegio y de la vida gracias por acompañarme, impulsarme y ser incondicionales, para ustedes mis más sinceras palabras de agradecimiento, sin ustedes no estaría donde estoy.

Gracias al profesor James por dejarme tener la combinación perfecta entre cantidad de terrenos a la nieve a pasarlo bien y trabajo duro de escritorio, en verdad, muy agradecido de las oportunidades que se me han presentado.

A todos y a cada uno de ustedes que son parte de mi vida, muchas gracias.

Tabla de Contenido

Índice de Tablas

Índice de Ilustraciones

Capítulo 1

Introducción

En la región Metropolitana, la agricultura y el agua disponible para el uso de la población se ven fuertemente influenciados por la cantidad de nieve que se forma en la cordillera de los Andes durante la temporada invernal. Es por esto que el agua disponible de los deshielos es un elemento clave para el desarrollo de la región, razón por la cual se realiza un estudio de la cantidad de agua disponible en la nieve (SWE) para la cuenca Piuquenes ubicada en la región Metropolitana.

El SWE depende de dos variables: la densidad y la profundidad de la nieve en el espacio. La primera se caracteriza por poseer una variación espacial baja en comparación con la profundidad de nieve [10] y una variación temporal considerable, esto es, que varía su valor a medida que avanza la temporada[18] [24]. Por otro lado, la profundidad de la nieve posee una alta variación espacial debido a los distintos fenómenos que la afectan (erosión, transporte por viento, derretimiento, entre otros).

Debido a la complejidad de los procesos que gobiernan la distribución espacial de la nieve, y además de la poca información meteorológica disponible en las zonas con aporte nival, es que realizar una modelación determinística de la profundidad de la nieve en el espacio se vuelve ilusorio. Por otro lado, para la realización de un modelo estadístico es necesaria la incorporación de datos que calibren el modelo. En este aspecto, uno de los mejores instrumentos para medir la profundidad de la nieve es el LiDAR debido a su gran precisión y densidad de datos que entrega en el espacio [8]. En el presente trabajo, se realiza una modelación de la variabilidad espacial de la profundidad de la nieve en la cuenca experimental Piuquenes con modelos de Machine Learning y Geoestadística utilizando datos LiDAR terrestre tomados en terreno durante el año 2018.

La variabilidad espacial de la profundidad de la nieve ha sido ampliamente estudiada mediante el uso de modelos estadísticos como árboles de regresión [3] [14] [25] [26], regresiones lineales [17] [23], k-Nearest neightboor [32]. Métodos de interpolación geoestadística como el kriging y co-kriging también han sido utilizado por diferentes autores [3] [13] [25], mostrando un gran potencial como interpolador.

Los modelos estadísticos de la profundidad de nieve utilizan como información de entrada

variables topográficas como la altura, pendiente, aspecto del terreno, entre otras. Winstral et al. (2002) [31] desarrollan en su estudio una variable denominada Upwind Slope (Sx). El Sx se utiliza para caracterizar la redistribución de la nieve por viento y queda definido en base a la topografía del lugar y la dirección preferencial del viento. Diferentes autores han documentado la importancia de este parámetro [13] [31] para cuencas en las cuales el viento juega un rol importante en la redistribución de la nieve por viento. Marofi et al (2011) [23] realizan una modelación de la profundidad de nieve en base a redes neuronales, en este trabajo los autores llegan a la conclusión que el factor Sx es el parámetro más influyente en los modelos propuestos.

Por otro lado, Revuelto et al (2014) [26] utilizan una variable llamada *Topographic Position Index (TPI)* que relaciona la altura de una celda de interés con la altura promedio en su entorno. Los autores utilizan árboles de regresión para modelar la profundidad de la nieve y concluyen que las variables más influyentes en los modelos son el TPI y Sx.

El principal objetivo de la memoria es la generación de modelos predictivos que puedan caracterizar la variabilidad espacial de la profundidad de la nieve en función de variables topográficas. En particular, se desarrollan dos tipos de modelos, de Machine Learning (ML) y otro geoestadístico en el cual se utiliza el método de co-kriging para captar la variabilidad espacial de la profundidad de nieve. Parte de los objetivos específicos de la memoria es caracterizar el grado de importancia de las variables independientes en los modelos, de forma tal de conocer cuál posee una mayor capacidad predictiva y de esta manera comprender de mejor manera los procesos que afectan a la distribución de la nieve en la cuenca Piuquenes.

Se espera con la memoria analizar la efectividad de los diferentes modelos propuestos mediante parámetros de desempeño típicos de los modelos de regresión, realizar un análisis de los errores de los modelos, intuir posibles fuentes de error y potenciales mejoras que se le podrían hacer.

El Capítulo 2 de la memoria muestra el Marco Teórico de los modelos en el cual se explican los diferentes modelos utilizados. Capítulo 3 muestra las variables topográficas (variables independientes) utilizadas en los modelos y además los escaneos de la profundidad de nieve para las 3 fechas que se tienen disponibles.

El Capítulo 4 muestra la metodología utilizada para llevar a cabo la modelación con ML y Geoestadística. El Capítulo 5 del estudio muestra los resultados obtenidos para los modelos propuestos.

Finalmente los capítulos 6 y 7 muestran el análisis de los resultados y conclusiones respectivamente, en estos capítulos se analizan los modelos planteados, posibles fuentes de errores, inferencias que entregan los modelos, etc.

Capítulo 2

Marco Teórico

La profundidad de la nieve es un parámetro que varía altamente en el espacio [22] [30] debido a que sobre el manto de nieve afectan diferentes procesos físicos, como el fenómeno de deposición, ablación, movimiento por avalanchas, sublimación, redistribución por viento, etc.

Diferentes autores han abordado el tema de la variabilidad espacial de la profundidad de nieve. Usualmente el problema es modelado a partir de modelos estadísticos como regresiones multivariadas [17] [23] y árboles de regresión [4] [25]. Los modelos de los autores utilizan variables topográficas como variables predictoras. Comúnmente la elevación del terreno, pendiente y orientación respecto al norte son las variables que se tienen en cuenta. A continuación se muestran dos variables usadas típicamente en los modelos de profundidad de nieve, Upwind Slope (Sx) y en menor medida, el Topographic Position Index (TPI).

Upwind Slope

Diferentes autores han incluido el Sx en sus modelos [25] [31] para caracterizar la influencia del viento en la distribución de la nieve. La variable muestra importancia en cuencas cuya distribución de la profundidad de nieve se encuentra controlada fuertemente por la redistribución de la nieve por viento.

Para determinar el valor de Sx se utiliza la ecuación 2.1. La ecuación incorpora la dirección del viento y el alcance en el cual se realizará el cálculo, ambos parámetros de la ecuación son utilizados para definir el pixel que determina la máxima pendiente en el terreno en la dirección del viento y con esto determinar el grado de exposición del pixel de interés.

$$Sx_{A,\mathrm{d}_{max}}(x_\mathrm{i}, y_\mathrm{i}) = max\left(tan^{-1}\left(\frac{H(x_v, y_v) - H(x_\mathrm{i}, y_\mathrm{i})}{[(x_v - x_\mathrm{i})^2 + (y_v - y_\mathrm{i})^2]^{0,5}}\right)\right) \tag{2.1}$$

donde A es el azimut definido por la dirección del viento, $(x_\mathrm{i}, y_\mathrm{i})$ son las coordenadas de la celda de interés y (x_v, y_v) son las coordenadas de las celdas ubicadas a lo largo del vector

de búsqueda definido por A y d_{max}.

El definir un solo azimut preferencial del viento es una tarea complicada, para evitar esto se definen diferentes azimuts y se toma el valor máximo de Sx.

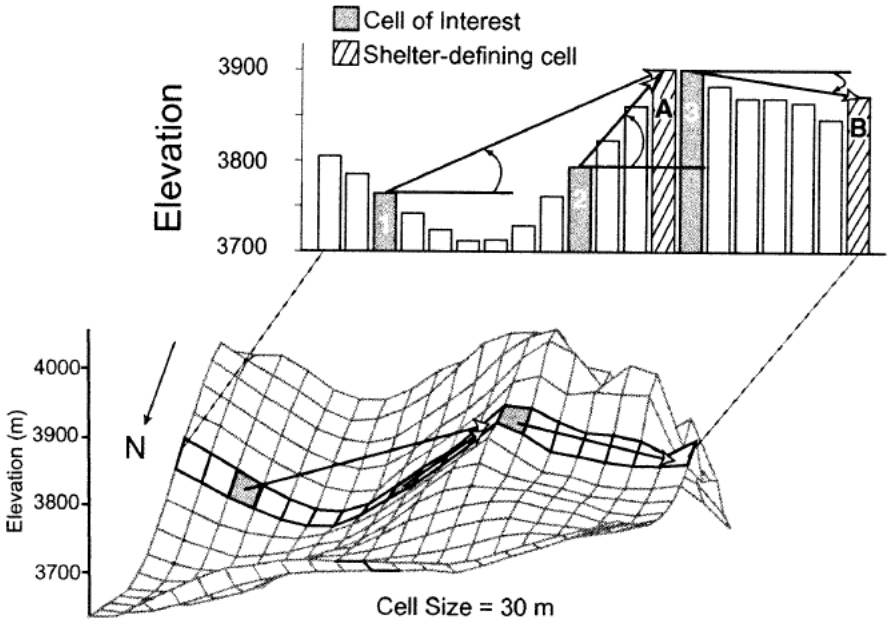


Figura 2.1: Esquema cálculo de Sx, fuente: Windstral et al (2002)

La figura 2.1 muestra un esquema del cálculo que se lleva a cabo para determinar Sx, en la imagen se nota que valores negativos de Sx significarán una celda mas expuesta, mientras que valores positivos de Sx indicarán una celda más protegida del viento.

Topografic Position Index (TPI)

Revuelto et al (2014) [26] utilizan el TPI como variable independiente en sus modelos de regresiones lineales y arboles de regresión. El parámetro relaciona la posición relativa entre el pixel de interés y los que se encuentran a su alrededor según la ecuación 2.2:

$$TPI = z - \bar{z} \tag{2.2}$$

En el estudio, los autores muestran al TPI como la variable independiente de mayor importancia en la mayoría de los modelos de árboles binarios y regresiones lineales.

2.1. Modelos de Machine Learning (ML)

Machine Learning (ML) es una herramienta para realizar análisis y modelamiento de fenómenos a partir de una muestra de datos. La principal característica de ML es que su aprendi-

zaje es basado en datos empíricos y puede modelar un fenómeno sin conocer las reglas físicas que lo gobiernan. Para efectos del presente trabajo se prueban modelos de Artificial Neural Netwroks (ANN), Random Forest (RF) y Gradient Boost Regresion (GBR). A continuación se muestra una breve explicación de los modelos utilizados junto a los principales parámetros que deben ser determinados para la modelación.

2.1.1. Modelo ANN

Existe una vasta documentación en cuanto a las diferentes partes que conforman un modelo de ANN, parte de los objetivos específicos de la memoria es lograr determinar la topología de la red que de el mejor resultado.

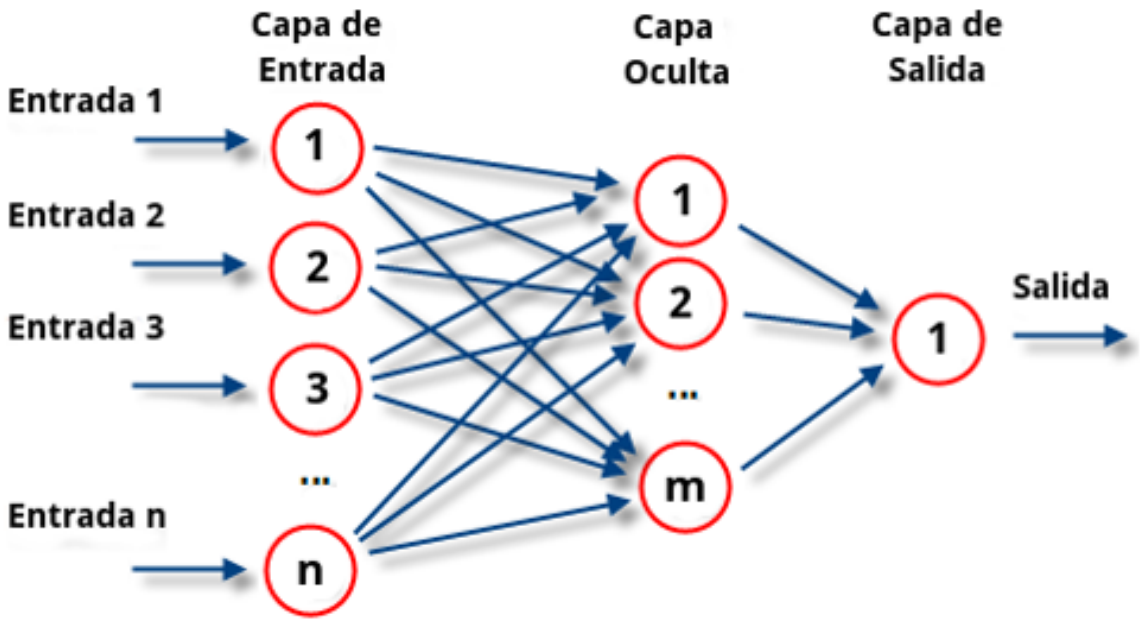


Figura 2.2: Partes de una red neuronal

En la figura 2.2 se muestra un esquema de red neuronal básico y se identifican 3 componentes:

1. **Capa de Entrada:** Se denomina capa de entrada a la parte de la red que recibe los inputs del modelo. Para la memoria esto corresponderá a las variables topográficas antes descritas. Estudios de ML muestran que se puede conseguir mejores resultados entregando además información geoestadística y análisis de errores en la capa de entrada [19].
2. **Capa Oculta (Hidden Layer):** En esta capa es donde se desarrollan los algoritmos de búsqueda y se genera el modelo. Esta capa es considerada muchas veces como una caja negra en la cual a través de iteraciones se va alcanzando la función objetivo.
3. **Capa de Salida:** Capa en la cual se encuentra el output del modelo, para el caso de la memoria, esto corresponderá a la profundidad de nieve.

Para llevar a cabo el aprendizaje, se deben definir diferentes variables, la notación y ecuaciones en este apartado están en base al libro de Kanevski et al (2009) [19].

Para entender el funcionamiento de una red neuronal con aprendizaje supervisado es necesario entender que sucede a nivel de neurona. En la figura 2.3 se identifican 3 partes fundamentales de una neurona:

- Conecciones a la neurona que están caracterizadas por los pesos w_{i}
- Una constante b_{i} la cual es una incógnita más en el problema de optimización
- Una función de activación o de transferencia $f(\cdot)$, la cual define la señal de salida de la neurona

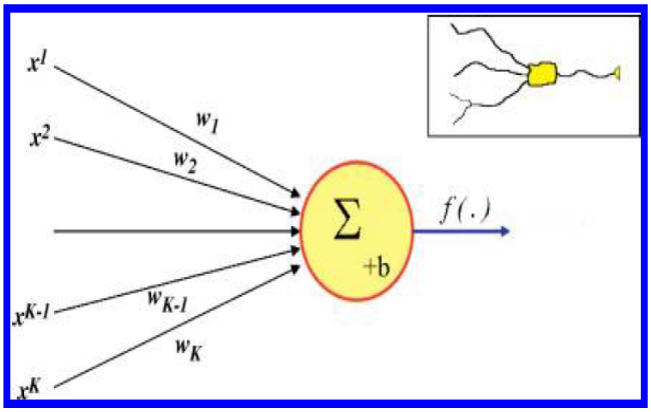


Figura 2.3: Representación proceso que ejecuta una neurona

La operación matemática que realiza una neurona cualquiera se muestra en la ecuación 2.3:

$$Z = f(\sum_{\mathrm{i}=1}^{K} w_{\mathrm{i}} x^{\mathrm{i}} + b) \tag{2.3}$$

Donde:

- $f(\cdot)$: Función de activación
- K: Número de neuronas
- w_{i}: Pesos
- x^{i}: Inputs que llegan a la neurona
- b: Constante del problema

La función de activación comúnmente son funciones con forma de s (sigmoidea). Las ecuaciones 2.4 y 2.5 muestran la función de activación logística y tangente hiperbólica respectivamente, ambas funciones son ampliamente usadas en modelos ANN:

$$f(x) = \frac{1}{1 + \mathrm{e}^{-x}} \tag{2.4}$$

$$f(x) = tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.5}$$

El problema de entrenar una neurona es un problema de optimización en el cual se desea determinar los valores de los w_i de forma tal que la señal de salida se asemeje al output deseado.

Para llevar a cabo la optimización de los pesos de la red, es necesario definir una función objetivo que se desea minimizar. La función que comúnmente se desea minimizar es el error de la red, este puede ser el RMSE (ecuación 2.23), MAE (ecuación 2.22), etc. Para efectos del trabajo se utilizará el MAE como función de error a minimizar.

2.1.2. Modelo Random Forest

El método de Random Forest es parte de una familia de modelos de ML denominados "Ensemble Models", que se basan en la generación de varios modelos base que se juntan para así formar una sola predicción. El método de RF utiliza como modelo base árboles de regresión cuyo funcionamiento se explica a continuación.

Árboles de Regresión

Los árboles de regresión funcionan en base a particionar el espacio de las variables independientes con el fin de obtener conjuntos más homogéneos en la variable de interés. La figura 2.4 muestra un árbol de regresión hecho en el trabajo de Winstral et al (2002) [31].

En la Figura 2.4 se pueden ver las diferentes particiones que se le hacen a la variable de interés en función de las variables independientes, además, cada nodo muestra la profundidad de nieve promedio (en metros) obtenida tras la división el set de datos. Los parámetros principales que se deben determinar para la modelación con árboles de regresión son la cantidad de nodos terminales del árbol y cantidad de datos que debe tener como mínimo un nodo.

Respecto a la cantidad de datos que debe tener un nodo, un número bajo de datos para nodos terminales disminuirá el error de entrenamiento pero funcionará mal para datos nuevos [6]. Se debe encontrar el equilibrio entre el número de nodos, cantidad de ramificaciones de forma tal de encontrar el modelo óptimo que ajuste los datos de manera adecuada.

El modelo de RF se caracteriza por utilizar el método de "Bagging", el cual es un algoritmo que genera subconjuntos de las variables de interés. Los subconjuntos son determinados mediante un muestreo aleatorio con reposición, esto es, que un mismo dato puede ser seleccionado mas de una vez y otros pueden no ser seleccionados [5][6].

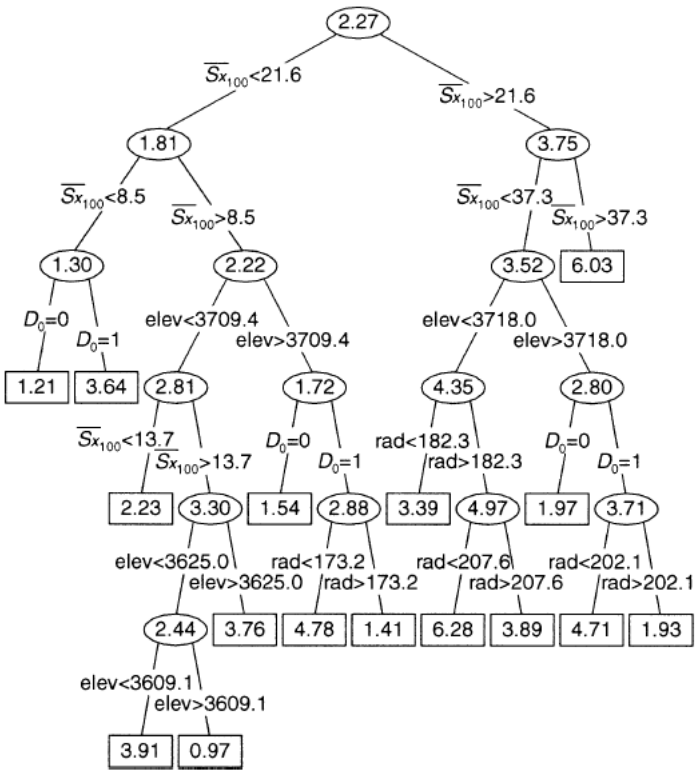


Figura 2.4: Árbol de regresión de la profundidad de nieve con 16 nodos terminales, las divisiones (split) se hacen en función de las variables independientes utilizadas por los autores. Los valores mostrados en los nodos corresponden a la profundidad de nieve promedio (m). Fuente: Winstral et al (2002)[31]

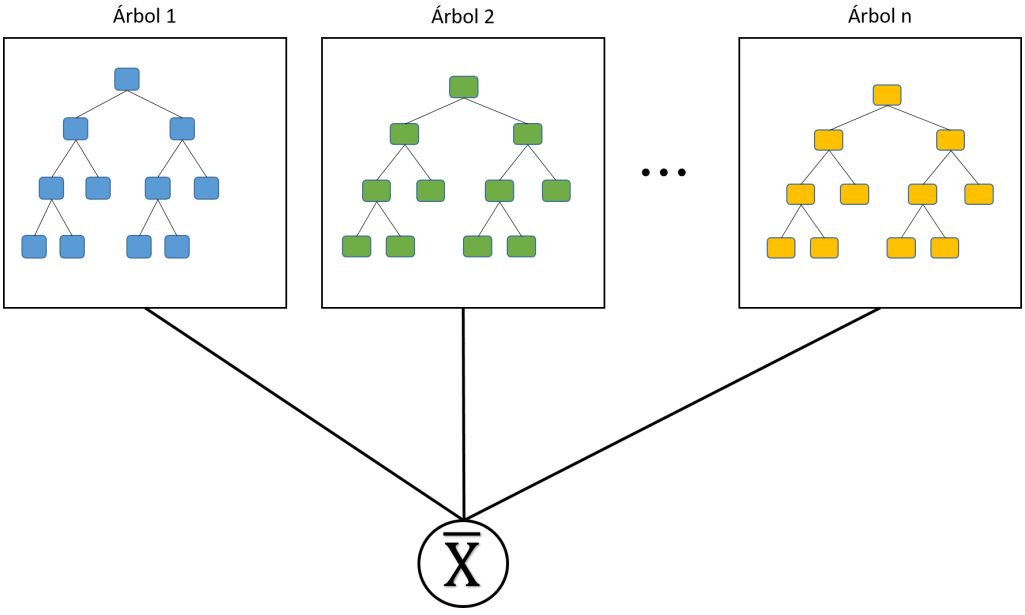


Figura 2.5: Fuente: Elaboración propia

El algoritmo de Bagging esta diseñado para mejorar la estabilidad y precisión de los modelos, además reduce la varianza y sirve para evitar un sobre-ajuste (overfitting) de los modelos [29], problema común a la hora de hacer modelos de aprendizaje supervisado.

Con los subconjuntos generados, se utilizan arboles de regresión (de ahí el nombre Random Forest) para generar predicciones, la predicción final sera un promedio de las predicciones individuales de cada árbol (ver Figura 2.5).

Los parámetros que deben ser determinados en RF son la cantidad de árboles que se utilizarán y la cantidad de datos mínimo que tendrá cada nodo, siendo el primer parámetro el de mayor influencia en la mayoría de los casos [5]. Además de los parámetros antes mencionados, también se determina la cantidad máxima de nodos, esto con el fin de controlar los recurso computacionales del modelo.

2.1.3. Modelo Gradient Boost Regressor (GBR)

Al igual que RF, GBR es parte de los denominados "Ensemble Models". El modelo GBR es un método iterativo en el cual se construyen modelos de regresión aditivos, esto es que el modelo final es la suma de todos los modelos base que se crean. GBR crea un árbol de regresión principal y luego forma árboles de regresión para trabajar los residuos del primer árbol [16].

De igual manera que en RF, se debe determinar las limitaciones de crecimiento de los árboles de regresión que se hagan, por lo que los parámetros a determinar son los mismos a los antes explicados en la sección de RF.

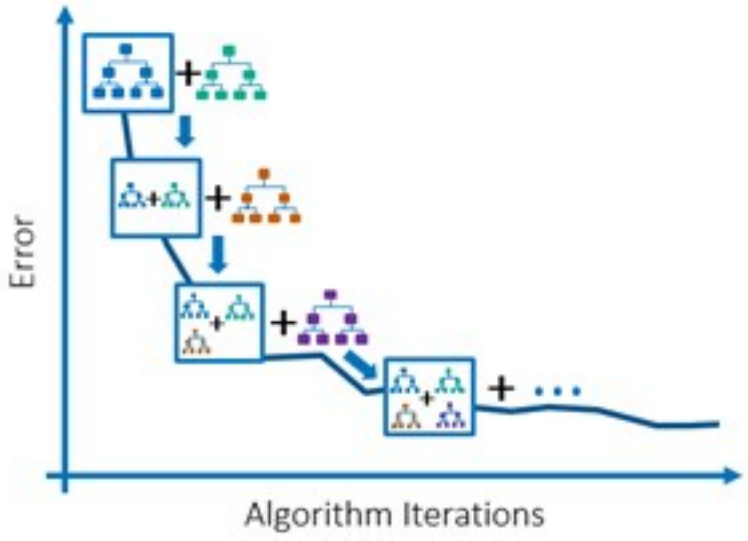


Figura 2.6: Ilustración funcionamiento de modelo GBR

2.2. Modelos Geoestadísticos

2.2.1. Bases de la geoestadística

La geoestadística se define como el estudio de fenómenos regionalizados, es decir, que se extienden en un espacio geográfico. Se intenta realizar una descripción matemática del fenómeno regionalizado a través de variables regionalizadas que miden ciertas propiedades o atributos del fenómeno. La extensión del estudio será el dominio del fenómeno regionalizado y se denomina campo.

La estadística clásica considera los datos como realizaciones independientes de una misma variable aleatoria, es decir, que siguen la misma distribución de probabilidad. Por otra parte, los modelos geoestadísticos consideran el valor z(x) de la medición en un punto x como una realización de una variable aleatoria Z(x). Esta hipótesis provoca que exista una familia de variables aleatorias, indexadas por las posiciones espaciales, las cuales constituyen una función aleatoria. Una característica vital de esta hipótesis es que las variables aleatorias no son independientes, por lo que existen correlaciones entre las mediciones. Las correlaciones que se encuentran entre las variables reflejan la continuidad espacial de la variable y esta a su vez se ve reflejada en el variograma.

2.2.2. Análisis Variográfico

Variograma Teórico

Se define el semi-variograma teórico (Ecuación 2.6) de una función aleatoria como sigue:

$$\gamma(x_1, x_2) = \frac{1}{2} var[Z(x_1) - Z(x_2)] \tag{2.6}$$

Comúnmente se omite el prefijo "semi" y se denomina simplemente variograma. Además, bajo el supuesto de estacionariedad de la función aleatoria, el variograma solo depende de la separación $(x_1 - x_2)$ entre los dos sitios considerados.

Variograma Experimental

Se define un estimador para el variograma teórico según la Ecuación 2.7:

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} [z(x_\alpha) - z(x_\beta)]^2 \tag{2.7}$$

Donde: α y β corresponden a coordenadas dentro del campo de estudio y $N(h)$ se define según la ecuación 2.8. La función $N(h)$ representa el numero de pares de datos que se encuentran separados a una distancia h. De esta manera, la ecuación 2.7 se puede interpretar

como el promedio de las diferencias cuadradas de pares de datos separados a una distancia h.

$$N(h) = \{(\alpha, \beta) \; tal \; que \; x_\alpha - x_\beta = h\} \tag{2.8}$$

Normalmente, si se definen separaciones exactas se tendrá un variograma con pocos puntos, razón por la cual se establece una tolerancia en la distancia de separación y en el azimut de búsqueda (Figura 2.7).

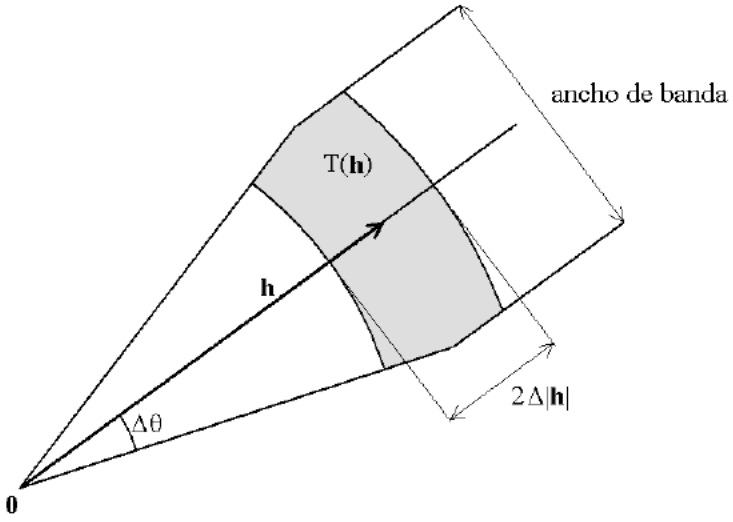


Figura 2.7: Región de tolerancia T(h), Fuente: Apunte de Geoestadística, Xavier Emery (2016).

Variograma Modelado

Es necesario ajustar un variograma modelado al variograma experimental de forma tal de poder representar la variable como una función continua.

Además se debe tener en cuenta que el variograma debe cumplir varias propiedades como paridad, positividad y debe ser una función de tipo negativo condicional [11] (ecuación 2.9). Esta última propiedad es una condición necesaria y suficiente para que la función sea un variograma. El adjetivo condicional indica que la desigualdad es valida sólo para ponderadores de suma total nula.

$$\forall k \in \mathbb{N}, \; \forall \lambda_1, ...\lambda_k \; tal \; que \; \sum_{\mathrm{i}=1}^{k} \lambda_\mathrm{i} = 0, \; \forall x_\mathrm{i}, ...x_k \in D, \sum_{\mathrm{i}=1}^{k} \sum_{\mathrm{i}=1}^{k} \lambda_\mathrm{i} \lambda_j \gamma(x_\mathrm{i} - x_j) \leq 0 \tag{2.9}$$

donde D es la extensión espacial de la zona de estudio.

Debido que la condición del tipo negativo condicional es una propiedad muy restrictiva y difícil de controlar, normalmente se utilizan funciones que se conoce cumplen con esta propiedad. En particular para el presente estudio se utilizan 3 funciones para modelar el variograma teórico: Efecto pepita, Modelo Esférico y Modelo Cúbico.

Modelo Efecto pepita:

$$\gamma(h) = \begin{cases} 0 & \text{si h=0} \\ C & \text{en caso contrario} \end{cases} \tag{2.10}$$

Donde C es la meseta del variograma que corresponde al valor en torno al cual el variograma se estabiliza. Para la modelación de los variogramas se considera que el variograma alcanza su meseta a partir de una distancia "a" denominado alcance del variograma.

Modelo Esférico:

$$\gamma(h) = \begin{cases} C(\frac{3|h|}{2a} - \frac{1}{2}(\frac{|h|}{a})^3) & \text{si } |\text{h}| < \text{a} \\ C & \text{en caso contrario} \end{cases} \tag{2.11}$$

Modelo Cúbico:

$$\gamma(h) = \begin{cases} C[7(\frac{h}{a})^2 - \frac{35}{4}(\frac{h}{a})^3 + \frac{7}{2}(\frac{h}{a})^5 - \frac{3}{4}(\frac{h}{a})^7] & \text{si } |\text{h}| < \text{a} \\ C & \text{en caso contrario} \end{cases} \tag{2.12}$$

El variograma experimental puede modelarse como una suma de varios variogramas modelados de. Se hace referencia a esta técnica como modelación del variograma mediante estructuras anidadas.

$$\gamma(h) = \gamma_1(h) + \gamma_2(h) + \gamma_3(h) + ...\gamma_s(h) \tag{2.13}$$

Caso multivariable

Las ecuaciones e inferencias antes mostradas corresponden al caso univariable. Para el caso multivariable, se hace diferencia entre variogramsa directos y cruzados como se muestra a continuación:

Sea Z_i y Z_j funciones aleatorias, se define le variograma cruzado para un vector h según la ecuación 2.14.

$$\gamma_{\text{i}j}(h) = \frac{1}{2} cov[Z_\text{i}(x+h) - Z_\text{i}(x), Z_j(x+h) - Z_j(x)] \tag{2.14}$$

El variograma simple o directo corresponderá al caso en que i $= j$ y en caso contrario se denomina como variograma cruzado.

El estimador del variograma cruzado se define de manera análoga al caso univariable pero con los subíndices i y j que indican las variables que están siendo consideradas en el variograma.

$$\hat{\gamma}_{\mathrm{i}j}(h) = \frac{1}{2|N_{\mathrm{i}j}(h)|} \sum_{N_{\mathrm{i}j}(h)} [z_{\mathrm{i}}(x_\alpha) - z_{\mathrm{i}}(x_\beta)][z_j(x_\alpha) - z_j(x_\beta)] \tag{2.15}$$

Notar que al ser $\mathrm{i} = j$ se recupera la ecuación 2.7.

2.2.3. Modelos de co-kriging

Co-kriging simple

El co-kriging simple (medias conocidas) es un método de predicción local de la variable de interés mediante las covarianzas explicadas por los variogramas.

Sea Z_1 la variable que se desea predecir en el sitio x_0 y m_{i} la media de la variable $Z_{\mathrm{i}}(\mathrm{i} = 1...N)$. El predictor se plantea como una combinación lineal de los datos de las N variables que se encuentren dentro de la vecindad:

$$Z_1(x_0)^* = a + \sum_{\mathrm{i}=1}^{N} \sum_{\alpha=1}^{n_{\mathrm{i}}} \lambda_\alpha^{\mathrm{i}} Z_{\mathrm{i}}(x_\alpha^{\mathrm{i}}) \tag{2.16}$$

donde a y los ponderadores $\lambda_\alpha^{\mathrm{i}}$ son las incógnitas del problema y siendo $\alpha = 1...n_{\mathrm{i}}$ correspondiente a las locaciones dentro de la vecindad del sitio x_0 para cada variable i.

Juntando la condición de esperanza nula del error (condición de insesgo o exactitud) y la condición de varianza mínima del error se llega al siguiente sistema de ecuaciones lineales.

$$\sum_{j=1}^{N} \sum_{\beta=1}^{n_j} \lambda_\beta^j C_{\mathrm{i}j}(x_\alpha^{\mathrm{i}} - x_\beta^j) = C_{\mathrm{i}1}(x_\alpha^{\mathrm{i}} - x_0) \ \forall \mathrm{i} = 1...N, \forall \alpha = 1...n_{\mathrm{i}} \tag{2.17}$$

El sistema de ecuaciones 2.17 resuelve el valor de los ponderadores de co-kriging, la ecuación 2.18 resuelve el valor de la constante a:

$$a = [1 - \sum_{\alpha=1}^{n_1} \lambda_\alpha^1] m_1 - \sum_{\mathrm{i}=2}^{N} [m_{\mathrm{i}} \sum_{\alpha=1}^{n_{\mathrm{i}}} \lambda_\alpha^{\mathrm{i}}] \tag{2.18}$$

Cabe destacar que las ecuaciones antes mostradas que definen el método de co-kriging son resueltas en una vecindad de x_0 que debe ser definida al momento de realizar la interpolación.

Co-kriging ordinario

Análogo al método de co-kriging simple, se plantea un sistema de ecuaciones utilizando las condiciones de linealidad, insesgo y varianza mínima del error pero ahora asumiendo desconocidas las medias de las variables estudiadas. A continuación se muestra el sistema de ecuaciones que se resuelve para este método:

$$\sum_{j=1}^{N}\sum_{\beta=1}^{n_j}\lambda_{\beta}^{j}C_{\mathrm{i}j}(x_{\alpha}^{\mathrm{i}}-x_{\beta}^{j})+\mu_{\mathrm{i}}=C_{\mathrm{i}1}(x_{\alpha}^{\mathrm{i}}-x_0)\ \forall \mathrm{i}=1...N,\forall\alpha=1...n_{\mathrm{i}} \tag{2.19}$$

$$\sum_{alpha=1}^{n_1}\lambda_{\alpha}^{1}=1 \tag{2.20}$$

$$\sum_{alpha=1}^{n_{\mathrm{i}}}\lambda_{\alpha}^{\mathrm{i}}=0\ \forall \mathrm{i}=2...N \tag{2.21}$$

En donde $\mu_1....\mu_N$ son incógnitas adicionales del problema (multiplicadores de Lagrange), se debe notar que el estimador de ambos tipos de co-kriging es el mismo y solo cambia el método por el cual son calculado los ponderadores.

2.3. Criterios estadísticos de los modelos

Existen diversos criterios para cuantificar los errores de los modelos. A continuación se muestran los errores que se utilizarán.

La ecuación 2.22 muestra el error absoluto promedio (MAE), esta métrica mide el promedio entre las distancias absolutas de la variable predicha y la medida.

$$MAE=\frac{\sum_{\mathrm{i}=1}^{n}|\hat{Z}_{\mathrm{i}}-Z_{\mathrm{i}}|}{n} \tag{2.22}$$

donde:

- $\hat{Z}$: Variable predicha.
- Z: Variable medida
- n: numero de datos

Otra métrica es la raíz del error cuadrado medio (RMSE) que se muestra en la ecuación 2.23. Tanto MAE como RMSE expresan el error promedio de los modelos, sin embargo, RMSE penalizará de mayor manera grandes errores. Por otra parte MAE posee la ventaja que el error promedio que entrega es más fácil de interpretar.

$$RMSE = \sqrt{\frac{1}{n}\sum_{\mathrm{i}=1}^{n}(\hat{Z}_{\mathrm{i}} - Z_{\mathrm{i}})^2} \tag{2.23}$$

Para analizar la bondad de ajuste de los modelos, se utiliza el coeficiente de determinacion o r^2 (ecuación 2.24).

$$r^2 = 1 - \frac{\sum_{\mathrm{i}}(\hat{Z}_{\mathrm{i}} - Z_{\mathrm{i}})^2}{\sum_{\mathrm{i}}(Z_{\mathrm{i}} - \bar{Z})^2} \tag{2.24}$$

Capítulo 3

Caracterización de la cuenca

3.1. Zona de estudio

La cuenca experimental Piuquenes se ubica en la comuna de Lo Barnechea, en la región Metropolitana. La cuenca experimental se encuentra inmersa en la hoya del río Mapocho con una altura que va desde los 3229 a los 4034 msnm. y una superficie de 2,6 km^2. La Figura 3.1 muestra los límites de la zona de estudio junto a las estaciones meteorológicas que miden variables como profundidad de nieve, temperatura del aire, presión atmosférica, humedad relativa, velocidad y dirección del viento.

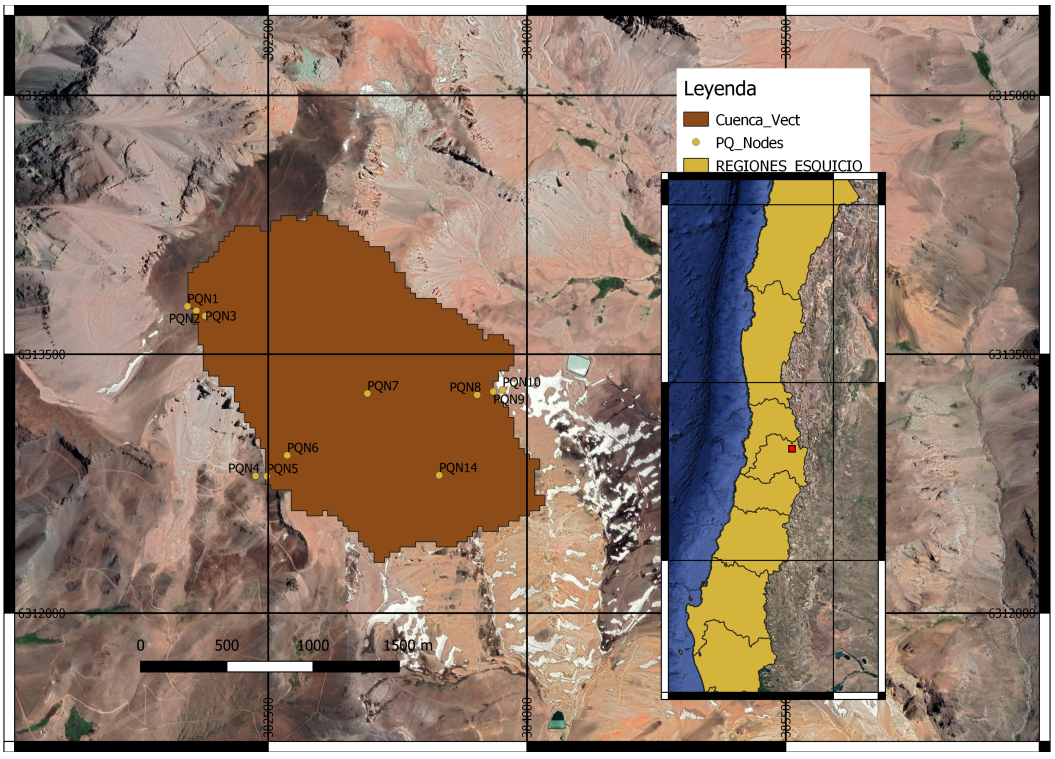


Figura 3.1: Delimitación cuenca Piuquenes

Para el estudio se efectuaron diferentes campañas de medición a lo largo del año 2018. La figura 3.2 muestra al LiDAR siendo utilizado en terreno. Se realizó una medición del terreno descubierto durante el verano (Figura 3.2a) y 3 mediciones de la profundidad de nieve en la época invernal del año 2018.

(a) Medición sin nieve (b) Medición con nieve

Figura 3.2: Fotos escaneo LiDAR en terreno

Para el estudio, se utilizan escaneos terrestres (TLS por su siglas en ingles). La Figura 3.2 muestra el escáner Reigl VZ-6000 LiDAR (Light Detection and Ranging) con el cual se generan las nubes de puntos que después se traducen en modelos digitales de elevación (DEM). Los DEM procesados poseen una resolución horizontal de 0.5 metros y un error en la vertical de $\sim$0.04 m.

3.2. Variables independientes

Se decide usar una resolución horizontal de cinco metros, por lo cual se realiza un reescalamiento de los productos LiDAR.

3.2.1. Elevación del Terreno (H)

Como primera variable independiente se tiene la altura topográfica (Figura 3.3). La imagen muestra un mapa digital de elevación (DEM) y es el producto de una medición con LiDAR terrestre tomada el 26 de enero del 2018, fecha en la cual se encontraba el terreno expuesto, libre de nieve.

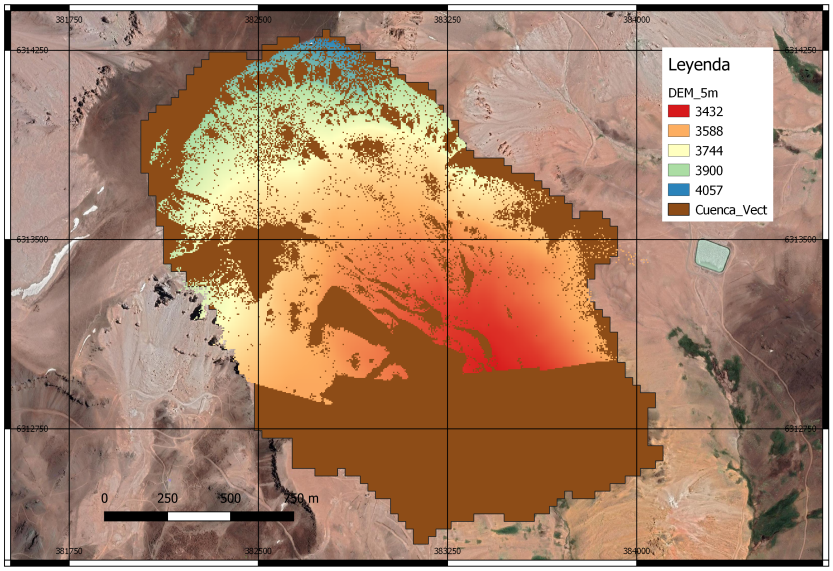


Figura 3.3: DEM LiDAR en la cuenca

En la Figura 3.4 se muestra el histograma para la elevación, en la imagen se aprecia que las mediciones efectuadas recorren un aproximado de 630 metros de desnivel partiendo desde los 3430 msnm. y llegando por sobre los 4000 msnm.

Tabla 3.1: Estadísticos Elevación

Variable	Min	Max	Rango	Promedio	Desvest
H	3432 m	4057 m	625 m	3650 m	130.1

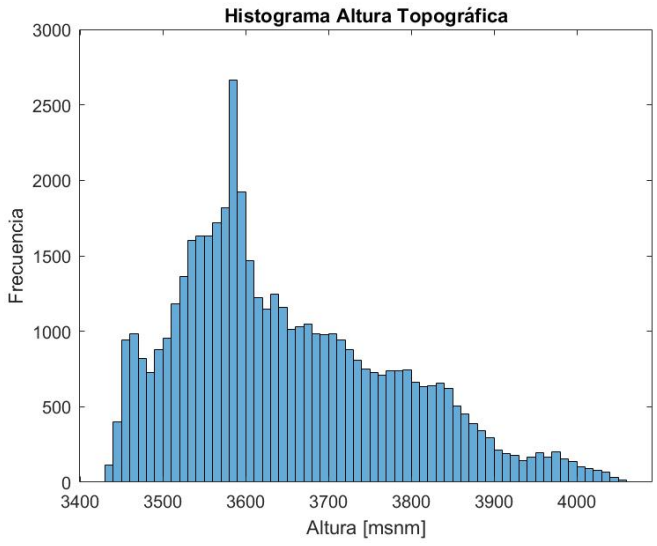


Figura 3.4: Histograma de la Elevación

3.2.2. Pendiente (Sl)

A partir del DEM de la cuenca (Figura 3.3), se calcula la pendiente del terreno utilizando la librería whitebox de python [21]. Este resultado se muestra en la Figura 3.5.

Se ilustran los datos mediante un histograma (Figura 3.6). La pendiente media de la cuenca medida es de 24.8° (Tabla 3.2) y la medición posee un rango de pendientes que va desde pendientes planas hasta pendientes escarpadas.

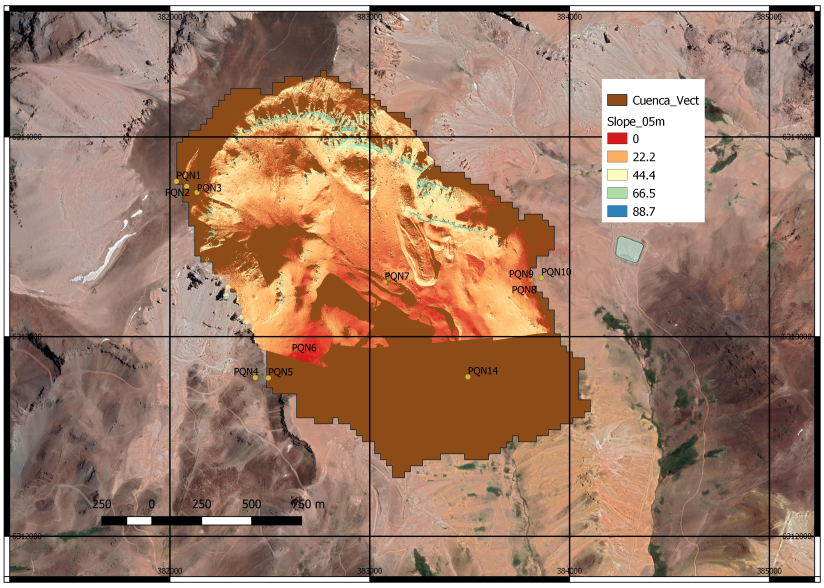


Figura 3.5: Pendiente de la cuenca

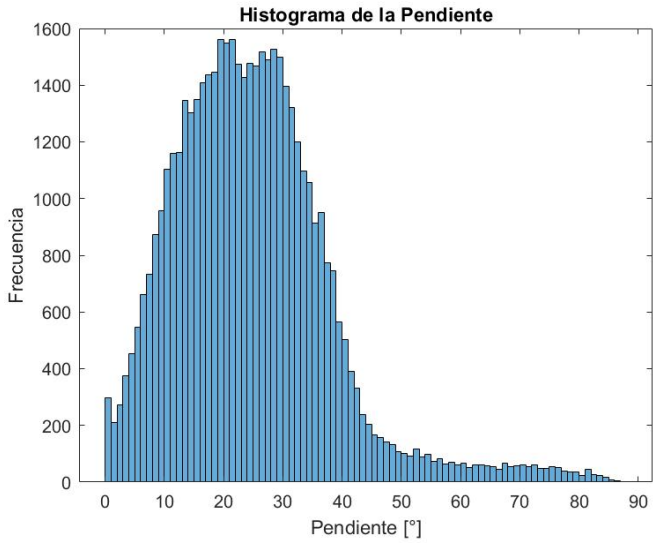


Figura 3.6: Histograma pendiente del terreno

Tabla 3.2: Estadísticos Pendiente

Variable	Min	Max	Rango	Promedio	Desvest
Slope	0°	87.3°	87.3°	24.8°	13.4°

3.2.3. Upwind Slope (Sx)

Se calcula el parámetro Sx según la ecuación 2.1. Sx se utiliza como una variable independiente que caracteriza el grado de exposición del terreno al viento.

En la Figura 3.7 se muestra un histograma de la dirección del viento medida en las estaciones meteorológicas instaladas en la cuenca. En la figura 3.7 se aprecia una dirección preferencial del viento con azimut 270°, correspondiente a vientos que provienen del oeste.

Se calcula Sx para un alcance de 100 metros [31]. La figura 3.8 muestra el máximo Sx calculado entre los 200° y 300° de azimut, conteniendo así el 68 % de los datos de dirección de viento. El calculo es hecho usando la librería whitebox [21].

Figura 3.7: Histograma de la dirección del viento

En la figura 3.9 se muestra el histograma de Sx. El histograma muestra que existen pocos valores negativos de la variable Sx. De acuerdo a lo explicado en el capítulo 2, esto quiere decir que se tienen pocos valores en los cuales se considere como terreno expuesto al viento.

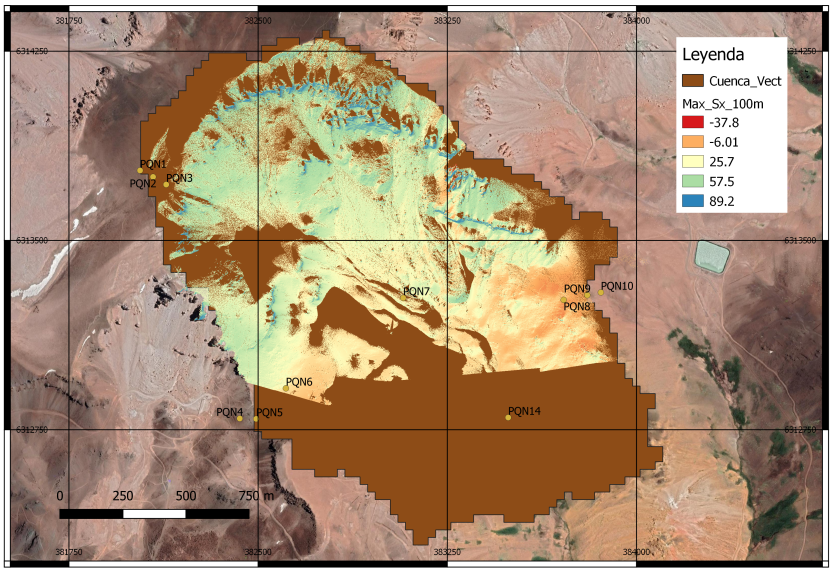


Figura 3.8: Máximo Upwind Slope para un alcance de 100 metros

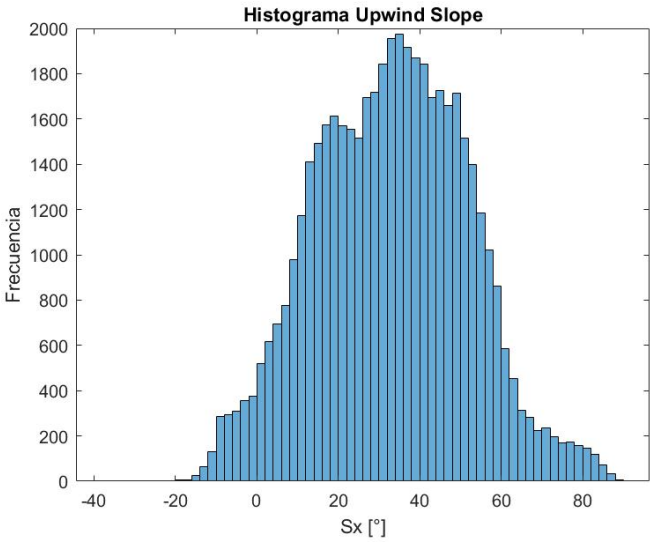


Figura 3.9: Histograma Upwind Slope calculado

Tabla 3.3: Estadísticos Upwind Slope

Variable	Min	Max	Rango	Promedio	Desvest
Sx	-37.4°	89.0°	126.4°	33.0°	18.8°

3.2.4. Curvatura (Cur)

Se obtiene la curvatura del terreno utilizando el comando *r.slope.aspect* de Grass 7.4.2. El programa entrega el resultado en unidades de m^{-1}, por ejemplo, una curvatura de 0.05 $m^{-}1$ corresponderá a un radio de curvatura de 20 metros. Según la convención, valores positivos de la curvatura indicaran un terreno convexo, mientras que valores negativos indican una curvatura cóncava.

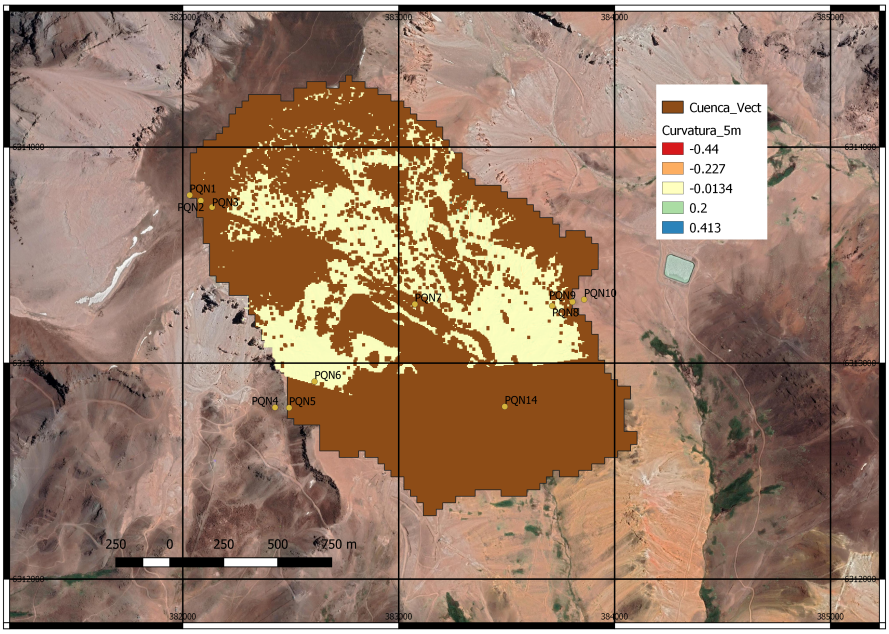


Figura 3.10: Mapa Curvatura

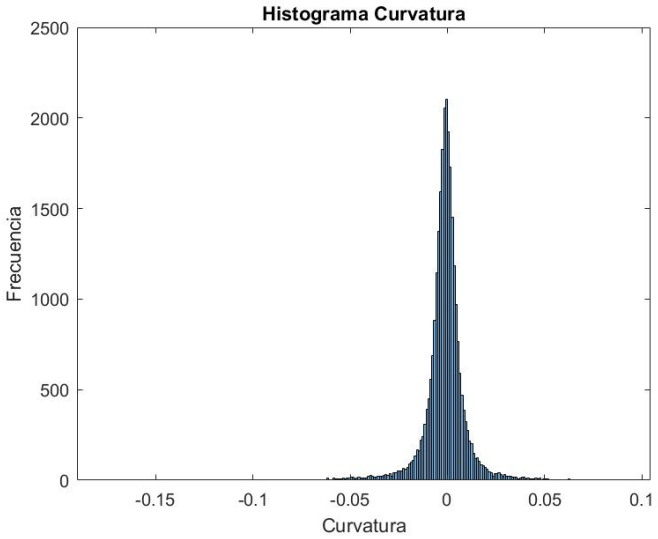


Figura 3.11: Histograma Curvatura del terreno

3.2.5. Topographic Position Index (TPI)

Usando la librería whitebox [21], se utiliza la función *diff from mean elev* para obtener el TPI.

Para el TPI se analizan diferentes alcances y se utiliza el coeficiente de correlación de Pearson para analizar las correlaciones entre cada TPI y las variables de interés (Tabla 3.4). La Tabla 3.4 muestra que el alcance de 100 metros muestra la mayor correlación en promedio, por lo tanto, se utiliza ese alcance (Figura 3.12).

Tabla 3.4: Coeficiente de correlación de Pearson entre los TPI para diferentes alcances y las 3 fechas de profundidad de nieve.

Variable	SD1	SD2	SD3	Promedio
TPI_25	-0.280	-0.230	-0.197	-0.235
TPI_50	-0.309	-0.312	-0.272	-0.298
TPI_75	-0.290	-0.350	-0.315	-0.318
TPI_100	-0.259	-0.370	-0.346	-0.325
TPI_200	-0.171	-0.372	-0.389	-0.311

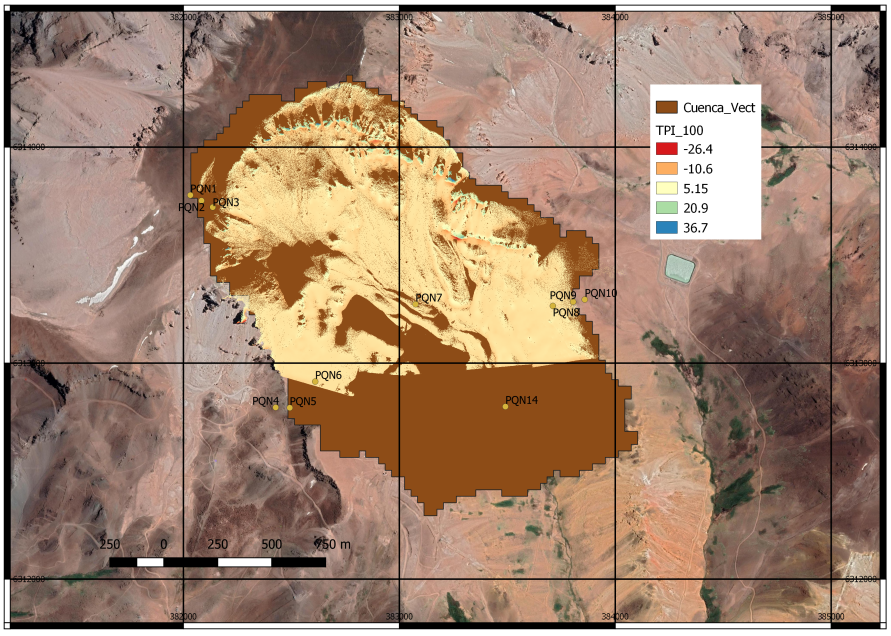


Figura 3.12: TPI para un alcance de 100 metros

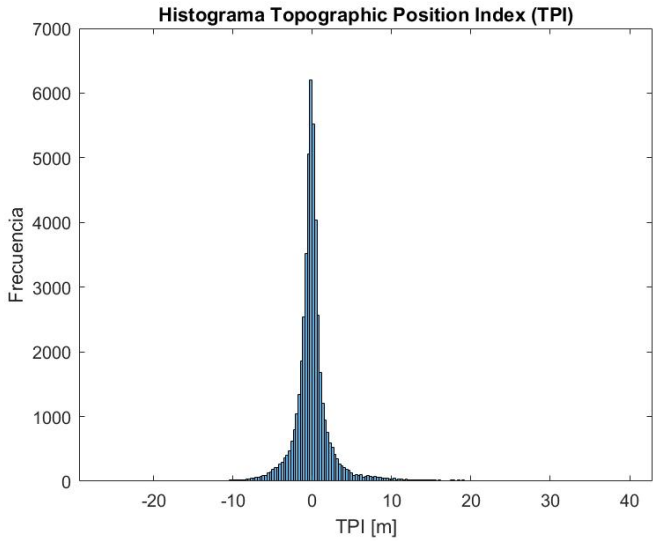


Figura 3.13: Histograma TPI de alcance 100 metros

3.3. Imágenes LiDAR de Profundidad de nieve

La Figura 3.14 muestra la precipitación medida en la estación Yerba Loca (curva azul) y las fechas de las mediciones LiDAR (lineas naranjas verticales) para la temporada invernal del 2018. En la imagen se aprecia que el primer y tercer escaneo (2 de junio y 9 de agosto respectivamente) fueron efectuados luego de eventos de tormenta. Además, se cuenta con la segunda medición (2 de agosto) efectuada previa a la tormenta del 7 de agosto.

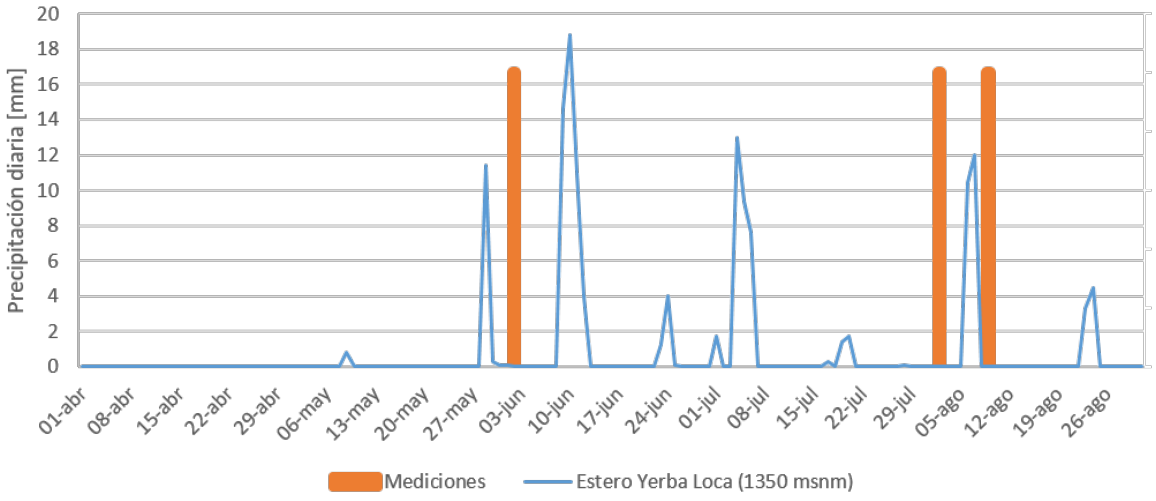


Figura 3.14: En azul, tormentas registradas en la estación Yerba Loca (1350 msnm) y en naranjo las fechas correspondientes a los escaneos de la profundidad de nieve.

LiDAR 02-06-2018 (SD1)

La Figura 3.15 muestra el producto del escaneo LiDAR efectuado el 2 de junio del 2018, la medición efectuada muestra una profundidad de nieve promedio de 24 centímetros (Tabla 3.5). Además la Figura 3.16 muestra la distribución de los datos mediante un histograma.

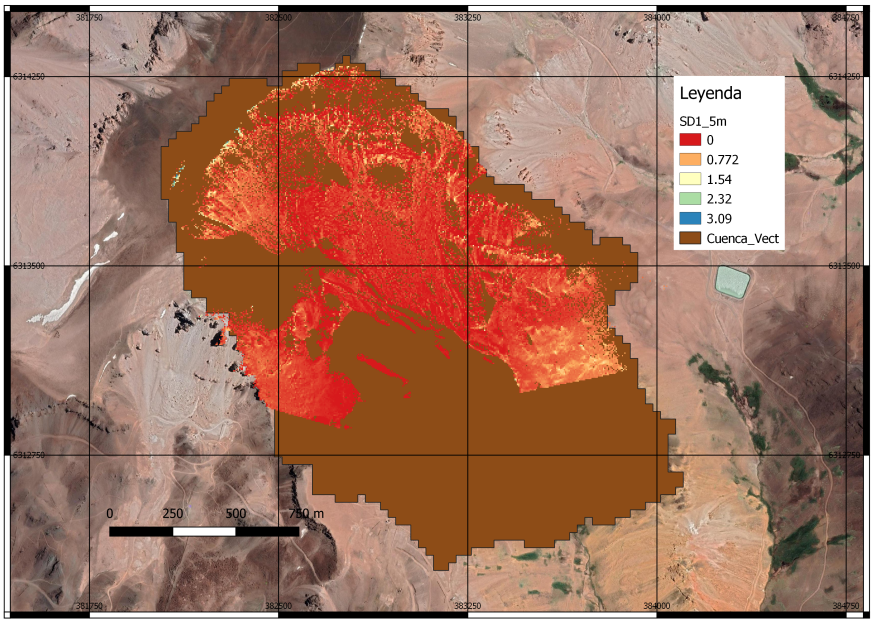


Figura 3.15: Profundidad de nieve medida el 02-06-2018

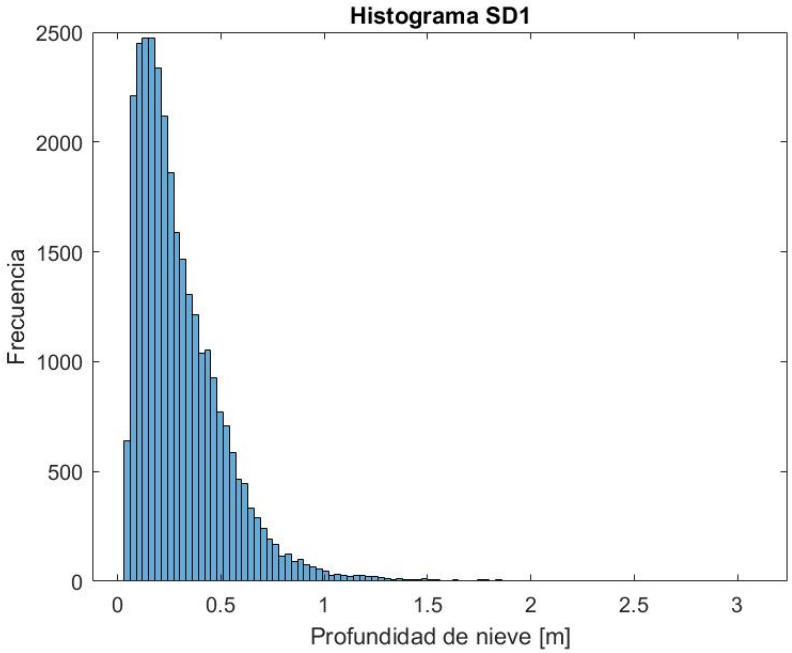


Figura 3.16: Histograma de la profundidad de nieve en la primera fecha de medición (SD1).

LIDAR 02-08-2018 (SD2)

La Figura 3.17 muestra el producto del escaneo LiDAR efectuado el 2 de agosto del 2018, la medición efectuada muestra una profundidad de nieve promedio de 62 centímetros (Tabla 3.5). Además la Figura 3.18 muestra la distribución de los datos mediante un histograma.

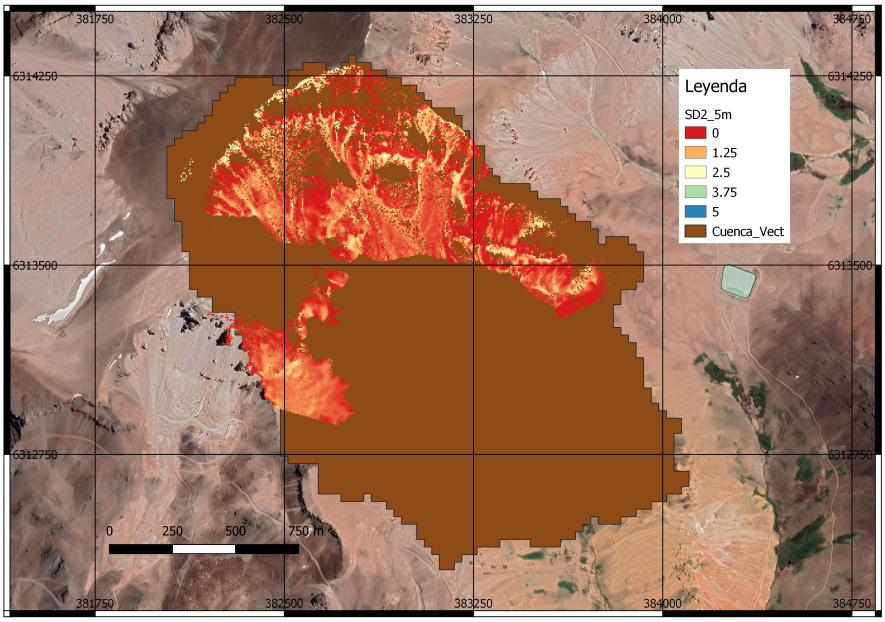


Figura 3.17: Profundidad de nieve medida el 02-08-2018

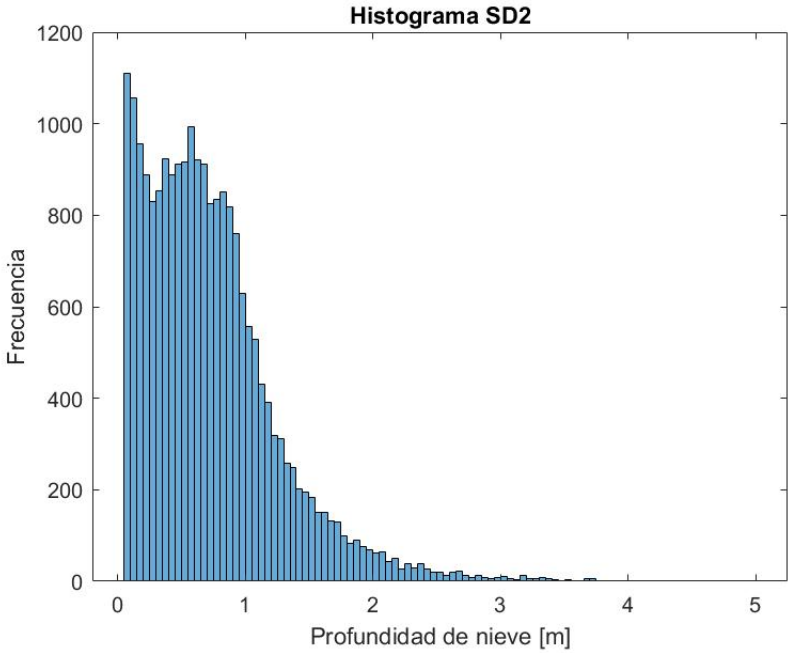


Figura 3.18: Histograma de la profundidad de nieve para la segunda fecha de medición (SD2).

LIDAR 09-08-2018 (SD3)

La Figura 3.19 muestra el producto del escaneo LiDAR efectuado el 9 de agosto del 2018, la medición efectuada muestra una profundidad de nieve promedio de 68 centímetros (Tabla 3.5). Además la Figura 3.20 muestra la distribución de los datos mediante un histograma.

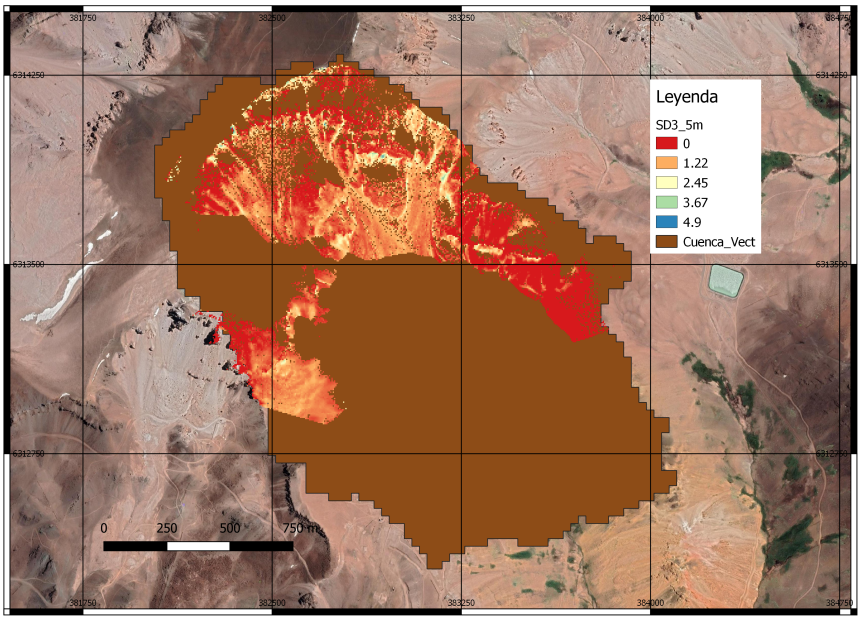


Figura 3.19: Profundidad de nieve medida el 09-08-2018

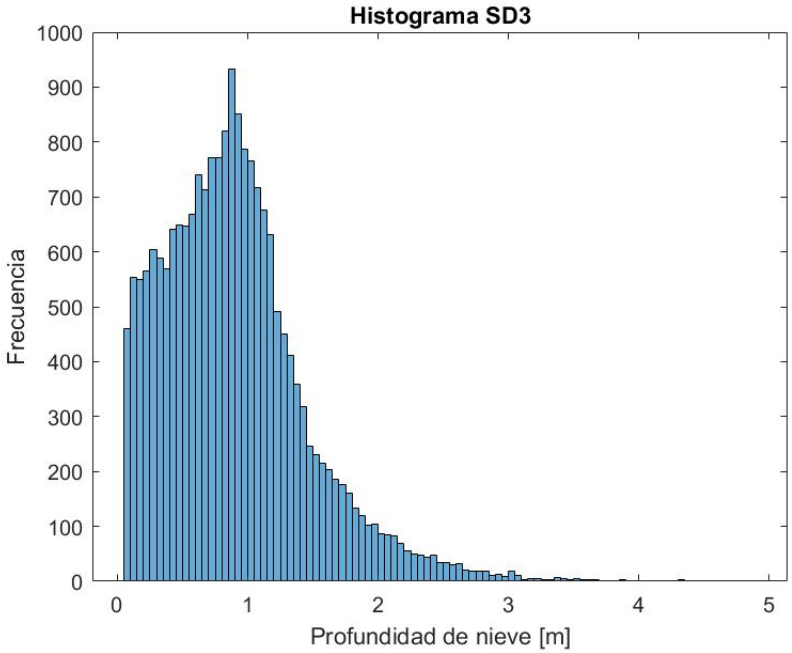


Figura 3.20: Histograma de la profundidad de nieve para la tercera fecha de medición (SD3).

La Tabla 3.5 muestra un resumen de los estadísticos más importantes de las variables mostradas en el capítulo.

Tabla 3.5: Resumen estadísticos de los datos

Datos	min	max	promedio	desvest	cv
SD1 [m]	0	3.088	0.240	0.238	0.99
SD2 [m]	0	4.996	0.617	0.554	0.90
SD3 [m]	0	4.899	0.676	0.617	0.91
H [msnm]	3432	4057	3650	130.1	0.04
Slope [°]	0	87.3	24.8	13.35	0.54
Sx [rad]	-0.653	1.553	0.577	0.328	0.57
Cur	-0.177	0.091	-0.001	0.011	-9.08
TPI [m]	-25.8	39.6	0.099	3.02	30.57

La Tabla 3.6 muestra los coeficientes de correlación de Pearson con los cuales se puede ver la correlación cruzada de cada variable. De la tabla, se nota que para la primera fecha la variable independiente que muestra mayor correlación es la Curvatura. Por otro lado, para la fecha 2 y 3 el TPI es la variable que muestra mayor correlación.

Tabla 3.6: Coeficiente de correlación de Pearson entre las variables del estudio.

	SD1	SD2	SD3	H	Sl	Sx	Cur	TPI
SD1	1	0.632	0.402	0.346	0.352	0.350	-0.429	-0.259
SD2	0.632	1	0.815	0.094	0.138	0.207	-0.358	-0.370
SD3	0.402	0.815	1	0.034	0.022	0.095	-0.311	-0.346
H	0.346	0.094	0.034	1	0.529	0.545	-0.019	0.045
Sl	0.352	0.138	0.022	0.529	1	0.794	-0.244	-0.103
Sx	0.350	0.207	0.095	0.545	0.794	1	-0.245	-0.120
Cur	-0.429	-0.358	-0.311	-0.019	-0.244	-0.245	1	0.385
TPI	-0.259	-0.370	-0.346	0.045	-0.103	-0.120	0.385	1

Capítulo 4

Metodología

4.1. Modelos de Machine Learning

Los tres modelos de Machine Learning propuestos son resueltos de manera análoga utilizando la librería *sklearn* de Python. A continuación se muestran los pasos seguidos para llevar a cabo la modelación.

1. Preprocesar los datos: Solo se utilizan los pixeles en los cuales se contiene toda la información (Figura 4.1), por lo cual se deben filtrar los datos.

2. Se dividen los datos de manera homogénea, en un conjunto de entrenamiento y otro de testeo (Figura 4.2) dejando el 70 % de los datos en el conjunto de entrenamiento y 30 % en el conjunto de testeo.

3. Se define la función objetivo a minimizar (loss function), para decidir cual será la función objetivo, se prueba el desempeño de los modelos para cada métrica. En la prueba, se nota que utilizar MAE (Ecuación 2.22) o RMSE (ecuación 2.23) como función objetivo no afecta fuertemente el resultado final de los modelos de ML hechos (no mostrado). Finalmente se decide utilizar MAE como función objetivo a minimizar y se realizan los análisis en torno a esta métrica.

4. Se identifican los parámetros necesarios para cada modelo, en el caso del modelo ANN se debe definir la función de activación, número de neuronas y algoritmo de optimización. Por otra parte para RF y GBR se deben definir parámetros que controlen el crecimiento de los árboles de regresión que se formen (número de nodos terminales, cantidad de árboles, etc).

5. Para cada configuración de parámetros se hace una validación cruzada con el algoritmo k-fold (Figura 4.3). El método divide en 3 partes iguales los datos del conjunto de entrenamiento (fold) y realiza el entrenamiento con 2/3 de los datos. Con el ultimo tercio se analiza el desempeño del modelo (validación). Para cada configuración de parámetros, el algoritmo realiza 3 entrenamientos y 3 validaciones, finalmente se entrega el MAE promedio de las 3 validaciones.

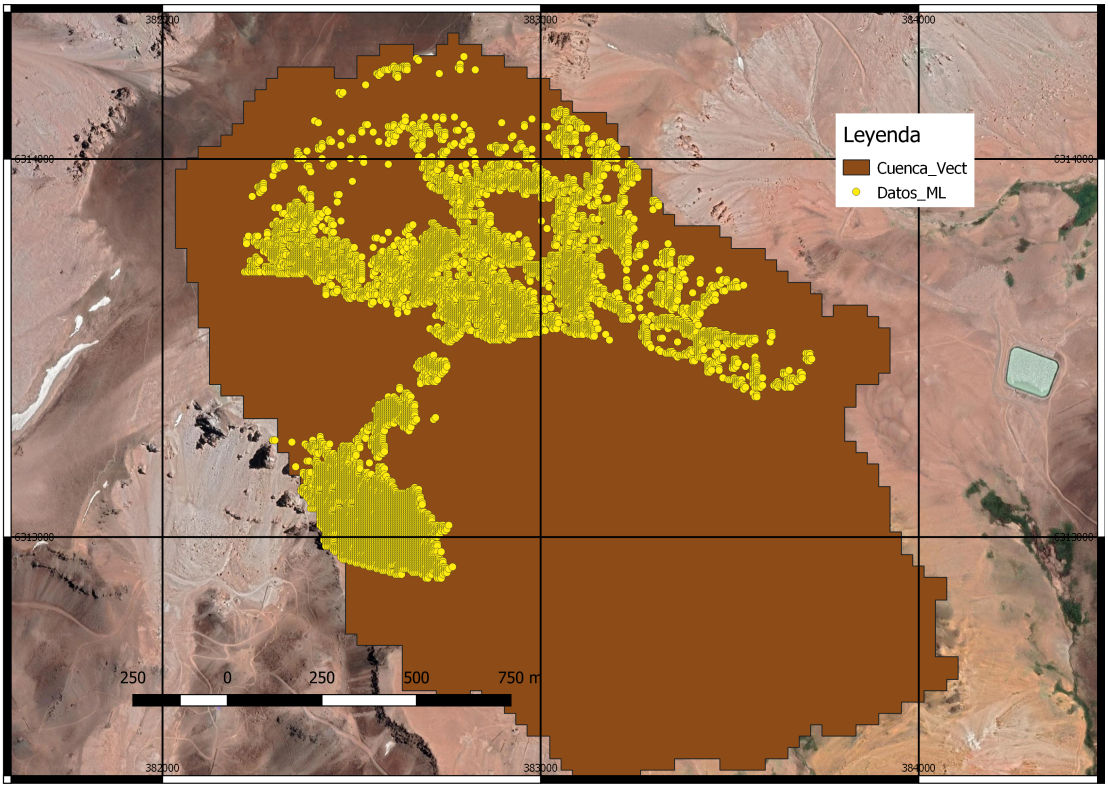


Figura 4.1: Distribución espacial datos para modelos de ML

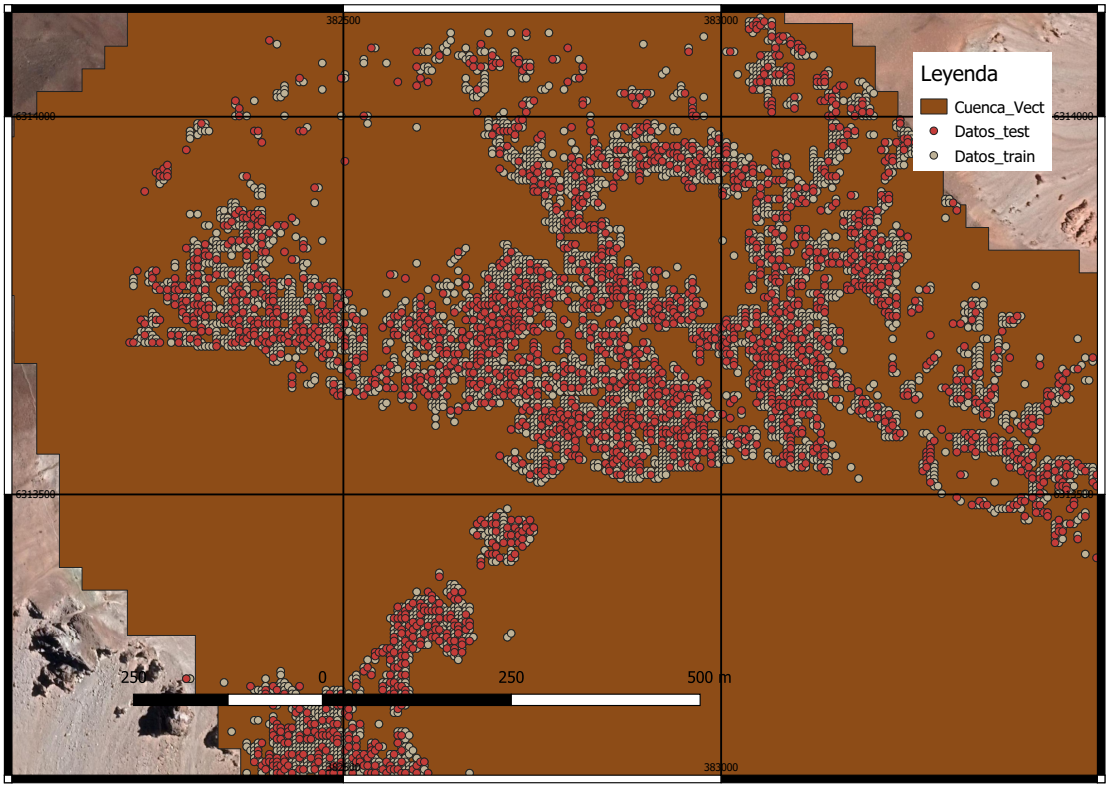


Figura 4.2: Datos de entrenamiento y testeo

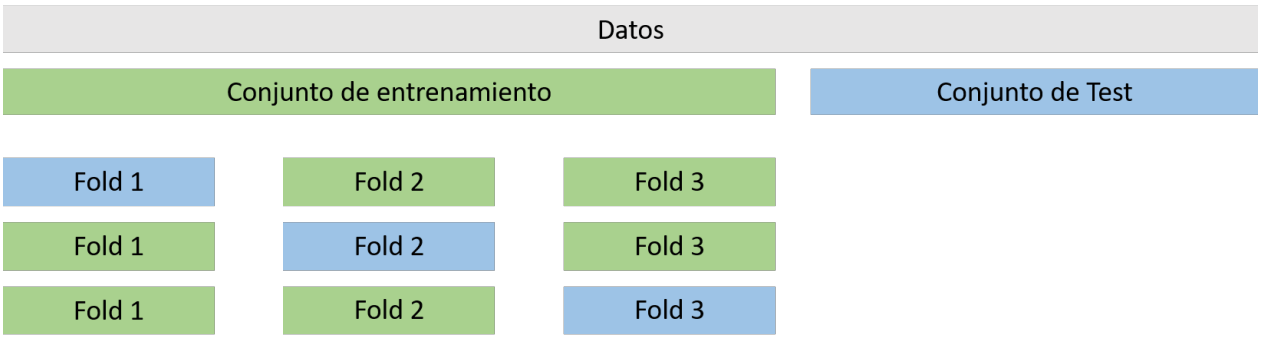


Figura 4.3: Validación cruzada, cajas verdes indican conjuntos que se utilizan para el entrenamiento mientras que las cajas azules se utilizan para analizar el desmpeño (validación).

6. De la validación cruzada, se analiza cual configuración de parámetros entrega el mejor desempeño en promedio, esta será la que entrega el menor MAE.
7. Se prueba el modelo con los datos de testeo y se analiza el desempeño obtenido.
8. Finalmente, con el fin de tener una idea de la influencia de las variables independientes en el modelo, se analizan los pesos y se caracteriza la importancia de cada variable para el modelo.

4.1.1. Modelo ANN

En el modelo ANN se analizan todas las combinaciones de los siguientes parámetros:

- Número de neuronas en la capa escondida: 100, 150, 200, 250, 300, 400, 500
- Función de activación: “tanh” y “relu”
- Learning Rate: Constante o Adaptable
- Learning Rate Inicial: 0.001, 0.01, 0.1
- Batch Size: 30, 50, 100, 120, 150, 200
- Máximo número de iteraciones: 100, 200, 300, 400, 500.

A partir de la validación cruzada, se seleccionan los parámetros que entregan el mejor desempeño en el conjunto de entrenamiento y se evalúa nuevamente el desempeño en el conjunto de testeo.

Finalmente se utiliza la librería de python *ELI5* para conocer la influencia de cada variable independiente en el modelo, siguiendo la metodología de *permutation importance* [1].

4.1.2. Modelo RF

Análogo al modelo ANN, se definen los rangos de los parámetros para analizar las diferentes combinaciones y así dar con el mejor modelo RF, a continuación se muestran los parámetros de búsqueda utilizados:

- Número de árboles de regresión: 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000
- Cantidad máxima de nodos terminales: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100
- Número de datos mínimo para nodo terminal: 2, 4, 8, 20, 50
- Bootstrap: True, False

Finalmente, al igual que en el modelo ANN se calcula la importancia de cada variable independiente en el modelo.

4.1.3. Modelo GBR

Para definir la topología del modelo de GBR se deben definir los mismos parámetros que para el modelo RF, es decir, parámetros para controlar el crecimiento de los arboles de regresión que son utilizados por el modelo:

- Número de árboles de regresión: 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000
- Cantidad máxima de nodos terminales: 20, 40, 60, 80, 100, 120, 140, 160, 180, 200
- Número de datos mínimo para nodo terminal: 2, 4, 8, 20, 50

4.2. Co-Kriging

A continuación se detalla la metodología empleada para llevar a cabo los modelos de co-kriging.

Variograma experimental

Para generar el modelo Geoestadístico se utilizan rutinas en MATLAB. La metodología sigue los pasos explicados en el Capítulo 2, donde primero se calcula el variograma experimental, luego se ajusta un variograma modelado, para finalmente resolver el sistema de ecuaciones dado por el co-kriging.

Se utiliza un código *gamv.m* para generar los variogramas experimentales, para utilizar esta función se definen los siguientes parámetros:

1. Dirección de búsqueda del variograma: Se modelan varias direcciones de búsqueda con el fin de ver si existen anisotropías en la profundidad de nieve. Finalmente, por simplicidad, se opta por utilizar un variograma omnidireccional.
2. Lag Separation: Se elige un lag de 10 metros con una tolerancia de 5 metros (Figura 2.7) y un total de 200 lag con lo que da un total de 2000 metros de variograma.

El output de esta rutina entrega los variogramas experimentales directos y cruzados de todas las variables del modelo.

Variograma Modelado

Se ajusta un variograma modelado a través de la función *vargfit.m* [11], la ventaja de esta rutina es que se modela el variograma de forma semi-automática debido a que solo es necesario definir los alcances de las estructuras anidadas y la varianza a priori de cada variable que se utilice.

Co-kriging

Para realizar el co-kriging se utiliza la función *cokrige.m*[12]. Para llevar a cabo el co-kriging se debe determinar la forma y tamaño de la vecindad de kriging, además se debe especificar el número de datos óptimo a tener en cuenta y el tipo de co-kriging que se desea hacer (simple u ordinario).

El modelo de co-kriging entrega la profundidad de nieve en el espacio y además la varianzas para cada punto. De esta forma, los resultados se ven resumidos en dos mapas, uno de profundidad de nieve predicha por el modelo y otro de las varianzas de los errores del modelo de co-kriging.

Para poder medir el desempeño del modelo se utiliza validación cruzada [9], con la metodología "Leave One Out" (LOO) con lo cual se obtienen los estadísticos de r^2 y MAE.

Capítulo 5

Resultados

5.1. Modelos de Machine Learning

Para cada medición LiDAR de la profundidad de nieve, se ajusta un modelo ANN, RF y GBR dando un total de 9 modelos de Machine Learning probados. Además, se realiza una regresión lineal multivariada (LR) con el fin de poder comparar los resultados de los modelos ML con un modelo sencillo como lo es una regresión lineal.

5.1.1. Modelo ANN

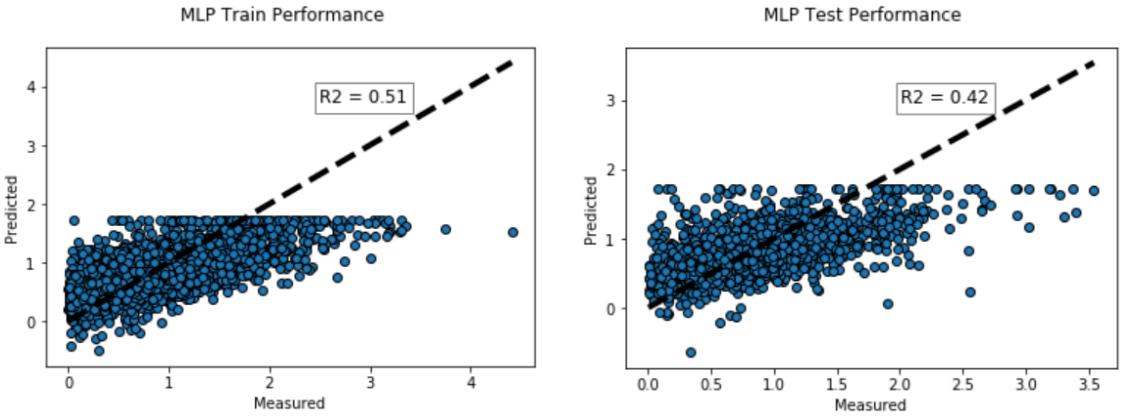
Se realiza la búsqueda del mejor predictor, pasando por todas las combinaciones de las variables mencionadas en el capítulo 4. La Tabla 5.1 muestra los parámetros que entregan el mejor desempeño en la validación cruzada:

Tabla 5.1: Parámetros de la red obtenidos mediante validación cruzada para cada fecha

Parámetros	SD1	SD2	SD3
Función de activación	relu	relu	relu
Batch size	150	150	300
hidden layer size	300	150	400
learning rate init	0.001	0.001	0.001
max iter	200	500	500

Luego se utilizan los parámetros de la Tabla 5.1 para analizar el desempeño del modelo ANN en el conjunto de testeo, la Figura 5.1 muestra el resultado del modelo para la fecha LiDAR del 2 de agosto.

Finalmente se calcula la importancia de las variables independientes del modelo mediante el método Permutation Importance[1]. La Tabla 5.2 muestra estos resultados y se nota que para la primera fecha Sx es la variable de mayor influencia con un 37 %. Para las otras dos



(a) Entrenamiento (b) Testeo

Figura 5.1: Resultados modelo de redes neuronales

fechas el TPI con alcance de 100 metros es la variable de mayor importancia, influyendo un 42 % en el modelo.

Tabla 5.2: Importancia relativa de las variables independientes en el modelo

Feature	SD1	SD2	SD3
H	30 %	18 %	22 %
Sl	19 %	16 %	22 %
Sx	37 %	14 %	6 %
Cur	10 %	9 %	8 %
TPI	4 %	42 %	42 %

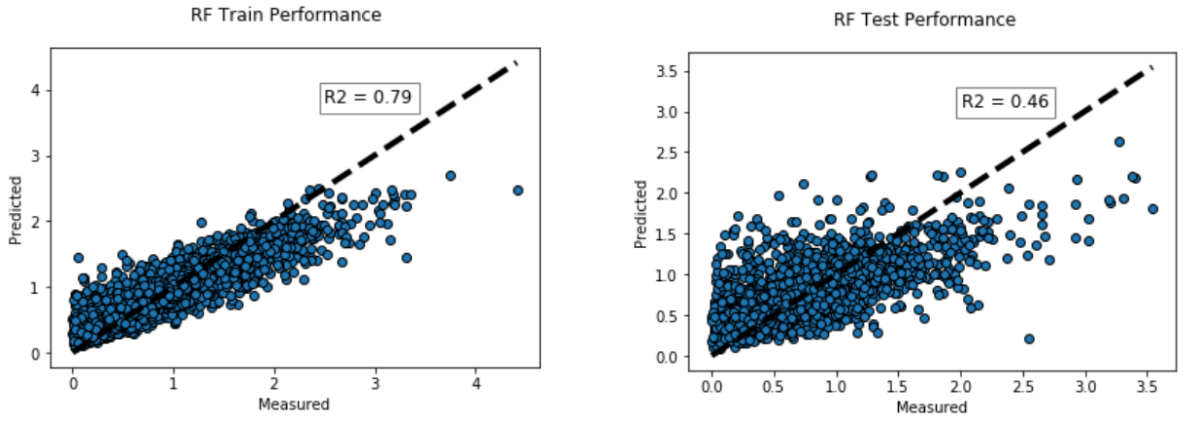
5.1.2. Modelo RF

Análogo al modelo ANN, se determinan los mejores parámetros obtenidos mediante validación cruzada para cada fecha LiDAR. Estos resultados se muestran en la tabla 5.3.

Tabla 5.3: Parámetros obtenidos de la validación cruzada para las 3 fechas LiDAR

Parámetros	SD1	SD2	SD3
Bootstrap	True	True	True
N° máximo de nodos	30	70	70
N° de datos nodo terminal	4	4	4
N° de datos para hacer división	25	5	5
N° de árboles de regresión	1600	1200	1200

Luego se utilizan los parámetros de la Tabla 5.3 para analizar el desempeño del modelo RF en el conjunto de testeo. La Figura 5.2 muestra el resultado del modelo para la fecha LiDAR del 2 de agosto.



(a) Entrenamiento (b) Testeo

Figura 5.2: Resultados Random Forest

Hecha la modelación, se puede calcular la importancia de los parámetros mediante el análisis de los pesos optimizados. La Tabla 5.4 muestra la importancia relativa de cada variable independiente en el modelo, y se aprecia que la Curvatura es la variable independiente con mayor influencia en la primera fecha, aportando un 30 %. Para las otras dos fechas (SD2 y SD3), el TPI es la variable independiente de mayor importancia en los modelos.

Tabla 5.4: Importancia relativa de las variables independientes en el modelo

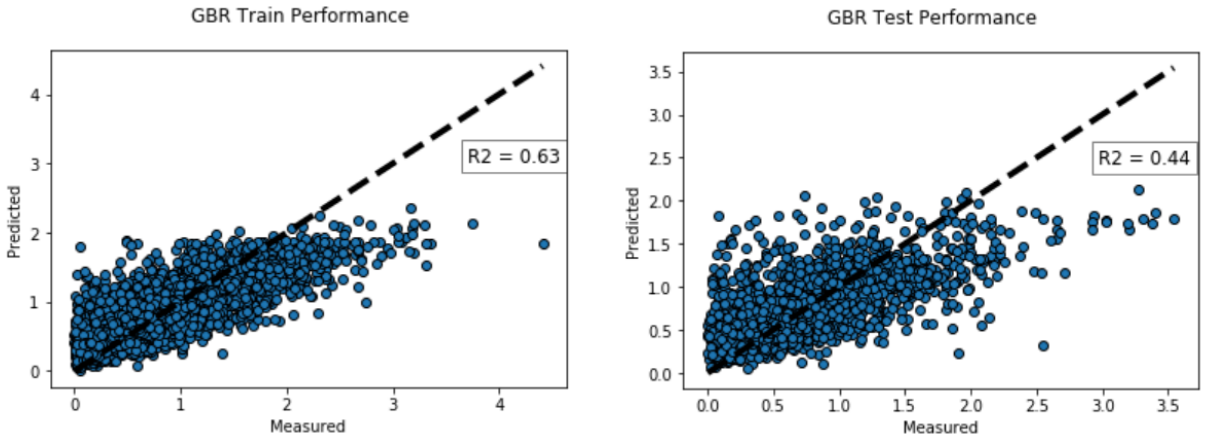
Features	SD1	SD2	SD3
H	25 %	21 %	29 %
Slope	18 %	13 %	13 %
Sx	10 %	12 %	10 %
Cur	30 %	12 %	10 %
TPI	17 %	42 %	37 %

5.1.3. Modelo GBR

Tabla 5.5: Parámetros obtenidos de la validación cruzada para las 3 fechas LiDAR

Parámetros	SD1	SD2	SD3
N° máximo de nodos	340	20	120
N° de datos nodo terminal	50	50	20
N° de datos para hacer división	100	200	2
N° de arboles de regresión	50	70	40

Utilizando los parámetros de la Tabla 5.5 se analiza el desempeño del modelo GBR para el conjunto de testeo. La Figura 5.3 muestra el resultado del modelo para la fecha LiDAR del 2 de agosto.



(a) Entrenamiento (b) Testeo

Figura 5.3: Resultados Gradient Boosting Regressor.

De igual manera que el modelo RF, se calcula la importancia de cada variable independiente en el modelo GBR (Tabla 5.6). Se llega al mismo resultado que en los modelos de RF, para la primera fecha, la curvatura es la variable que muestra mayor importancia mientras que para la segunda y tercera fecha es el TPI.

Tabla 5.6: Importancia relativa de las variables en el modelo

Features	SD1	SD2	SD3
H	29 %	20 %	31 %
Slope	15 %	11 %	13 %
Sx	7 %	9 %	11 %
Cur	35 %	11 %	9 %
TPI	13 %	49 %	36 %

5.1.4. Resumen Modelos ML

En la Tabla 5.7 se muestra un resumen de los desempeños de los modelos de ML para el conjunto de testeo, además se incluye el desempeño del modelo de regresión lineal:

Tabla 5.7: Resumen de los modelos

	SD1		SD2		SD3	
Modelo	r2	MAE	r2	MAE	r2	MAE
ANN	0.37	0.11	0.42	0.25	0.43	0.26
RF	0.44	0.10	0.46	0.23	0.48	0.23
GBR	0.44	0.10	0.44	0.24	0.46	0.24
LR	0.34	0.11	0.22	0.29	0.18	0.32

Se compara gráficamente el r2 (Figura 5.4) y el MAE (Figura 5.5) de los modelos.

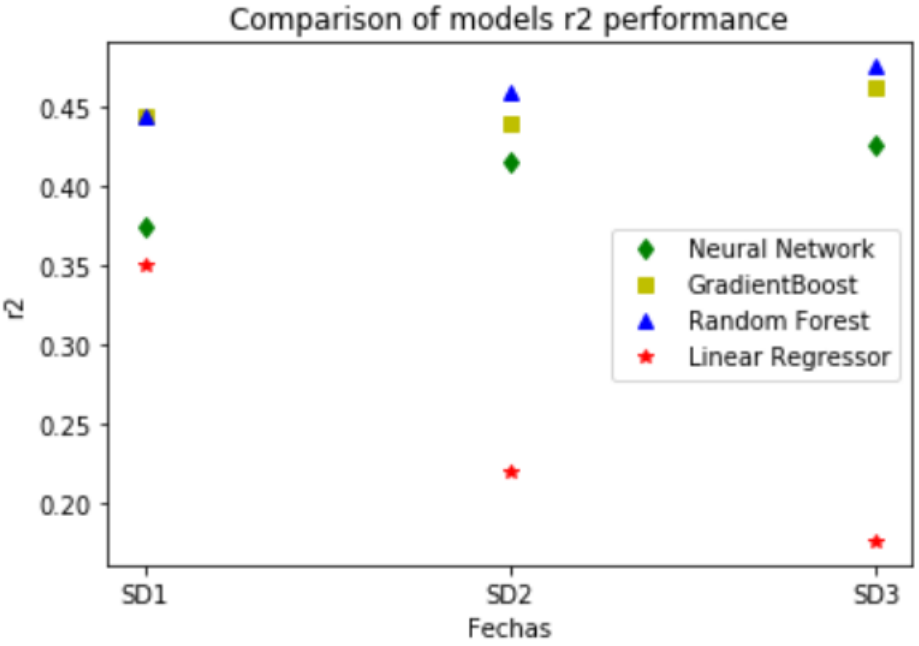


Figura 5.4: Comparación del r2 de los modelos propuestos

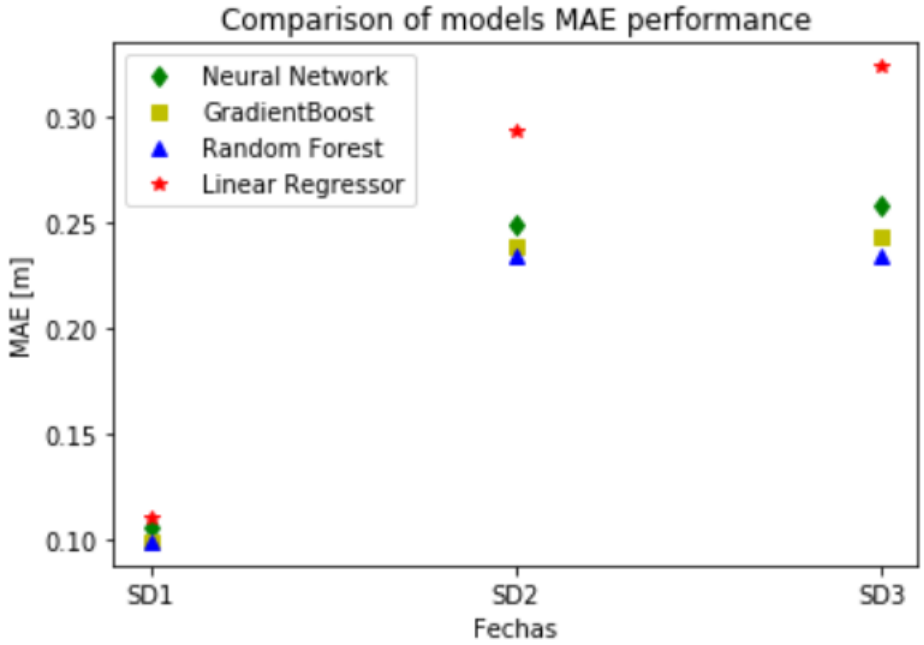


Figura 5.5: Comparación del MAE de los modelos propuestos

5.2. Modelo Geoestadístico

A continuación se muestran los resultados para el modelo geoestadístico. Los resultados expuestos a continuación consideran las mismas variables independientes utilizadas en los modelos de ML.

En la presente sección se muestran los resultados principales del modelo geoestadístico para la medición LiDAR del 2 de agosto.

5.2.1. Análisis variográfico

De acuerdo a la metodología expuesta en el capítulo 4, lo primero que se hace es analizar el variograma de la profundidad de nieve en 6 direcciones principales (Figura 5.6).

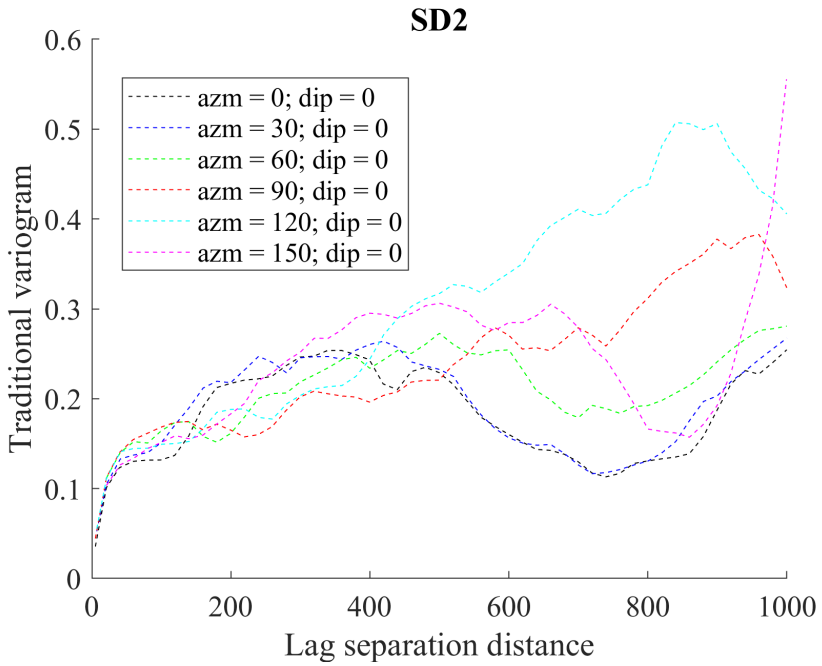


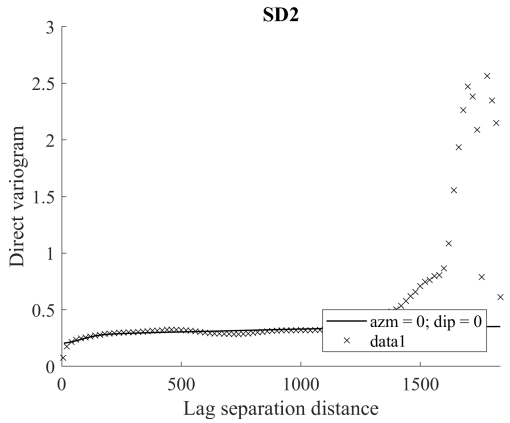
Figura 5.6: Variograma simple para 6 direcciones de la profundidad de nieve de la fecha 2.

De la Figura 5.6 se nota que existen anisotropías en la variable de interés, sin embargo, por simpleza, se opta por hacer un variograma omnidireccional. El trabajo del análisis de las anisotropías presentes en la profundidad de la nieve se escapa del estudio y queda propuesto como posible trabajo a futuro.

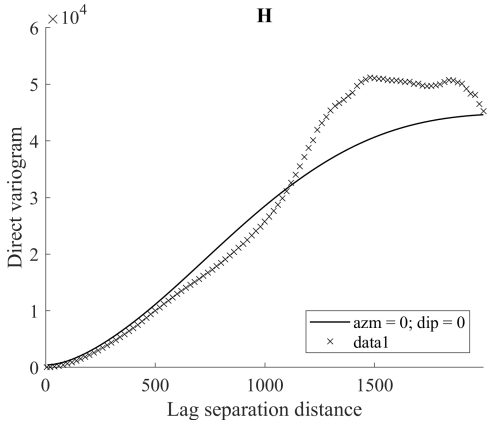
Se utilizan estructuras anidadas (Tabla 5.8) para modelar el variograma. La Figura 5.7 muestra los ajustes hechos de los variogramas simples para SD2.

Tabla 5.8: Modelos anidados utilizados para el ajuste de los variogramas

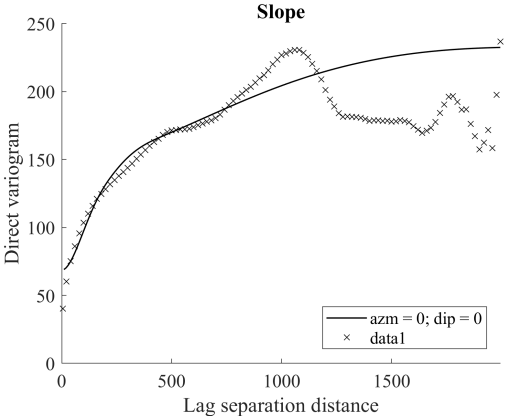
Modelo Anidado	Alcance [m]
Esférico	10
Cúbico	250
Cúbico	500
Cúbico	1500
Cúbico	2000



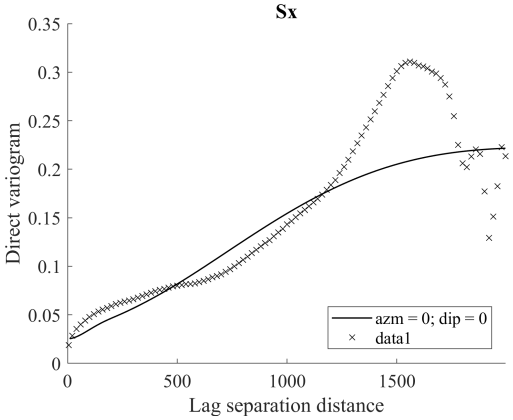
(a) SD2



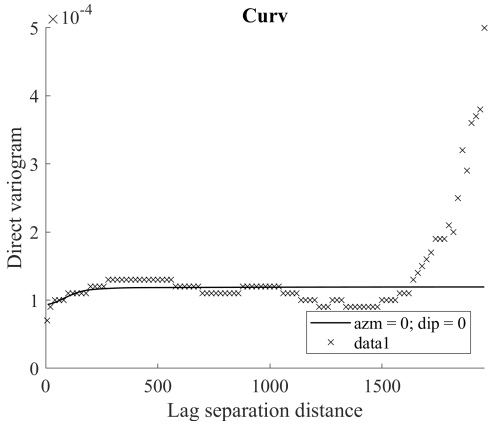
(b) H



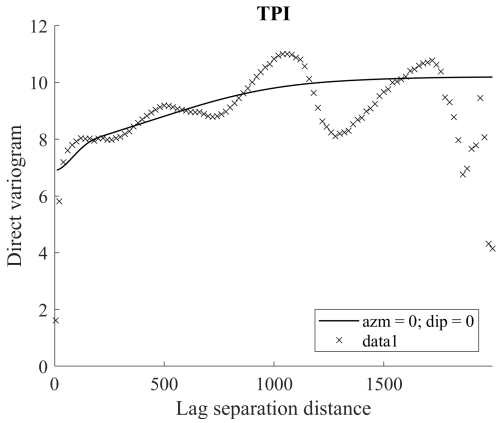
(c) Slope



(d) Sx



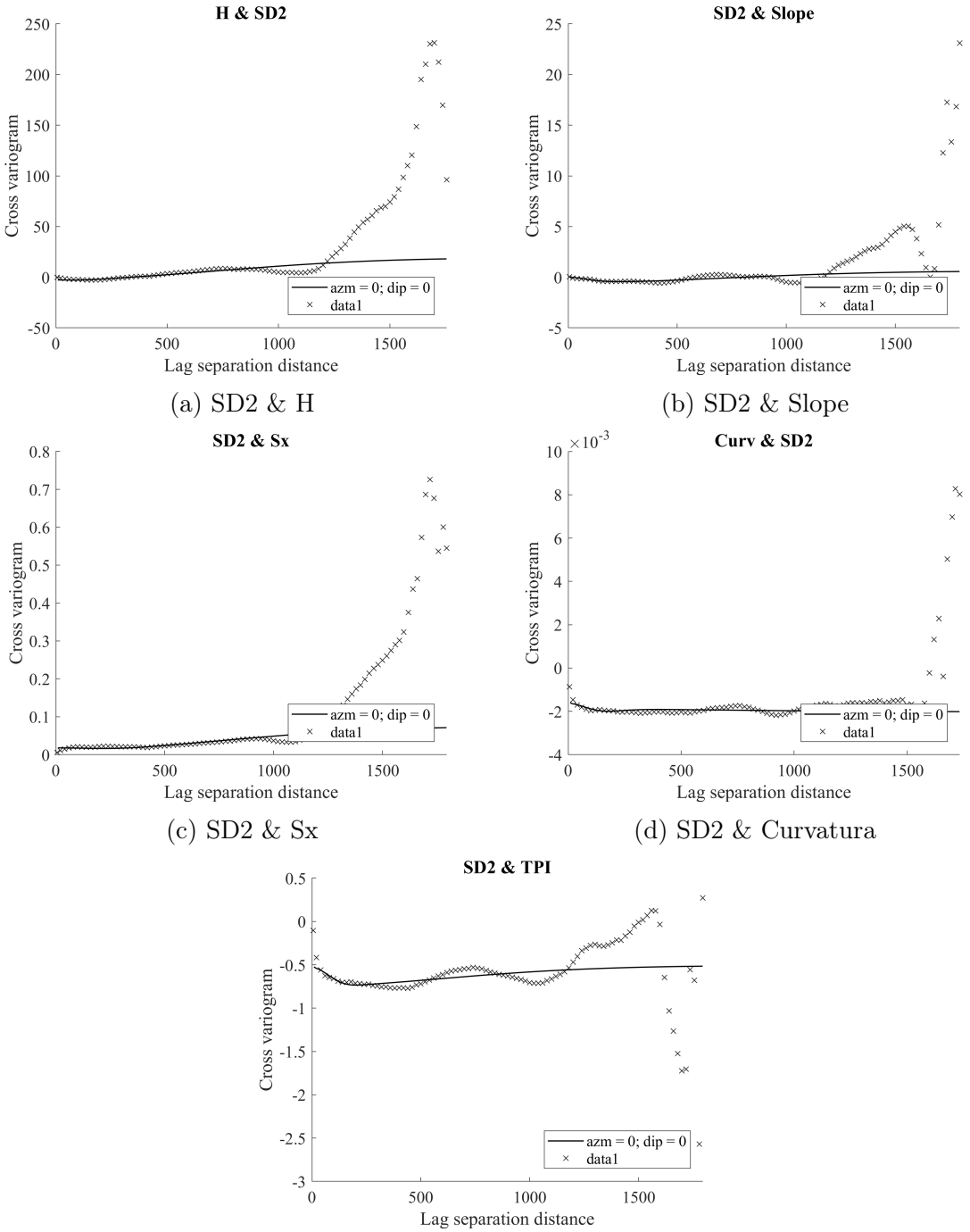
(e) Curv



(f) TPI

Figura 5.7: Variogramas directos modelados

Además se cuenta con los variogramas cruzados entre las variables independientes y la profundidad de nieve.



(a) SD2 & H

(b) SD2 & Slope

(c) SD2 & Sx

(d) SD2 & Curvatura

(e) SD2 & TPI

Figura 5.8: Variogramas Cruzados entre SD2 y las variables independientes utilizadas

5.2.2. Modelo de co-kriging

Con los variogramas directos y cruzados, se realiza el co-kriging. Se evalúa el desempeño de los modelos de co-kriging mediante validación cruzada, esto se muestra en la Tabla 5.9.

Tabla 5.9: Resultados modelos de Simple Co-Kriging (SCK) y Ordinary Co-Kriging (OCK)

Modelo	SD1		SD2		SD3	
	r2	MAE	r2	MAE	r2	MAE
SCK	0.48	0.370	0.76	0.161	0.48	0.373
OCK	0.48	0.373	0.76	0.183	0.48	0.373

Se obtiene el mapa de la distribución de la profundidad de nieve para el 2 de agosto:

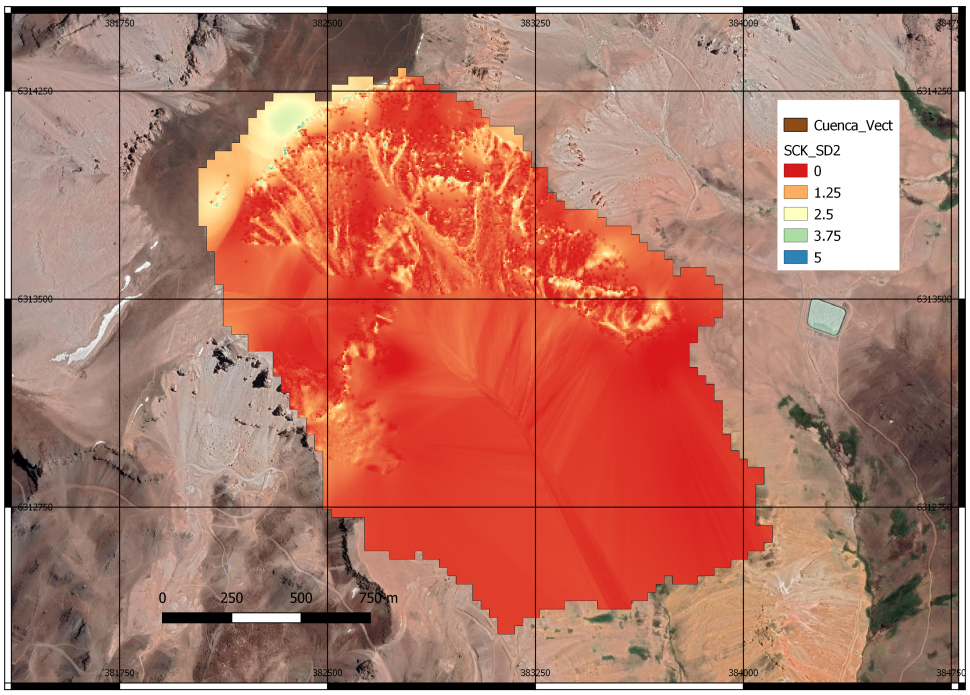


Figura 5.9: Profundidad de nieve para SD2 con el modelo SCK

Además, el modelo de SCK entrega un mapa de la varianza esperada del error de la predicción, la Figura 5.10 muestra el mapa de la varianza de la predicción.

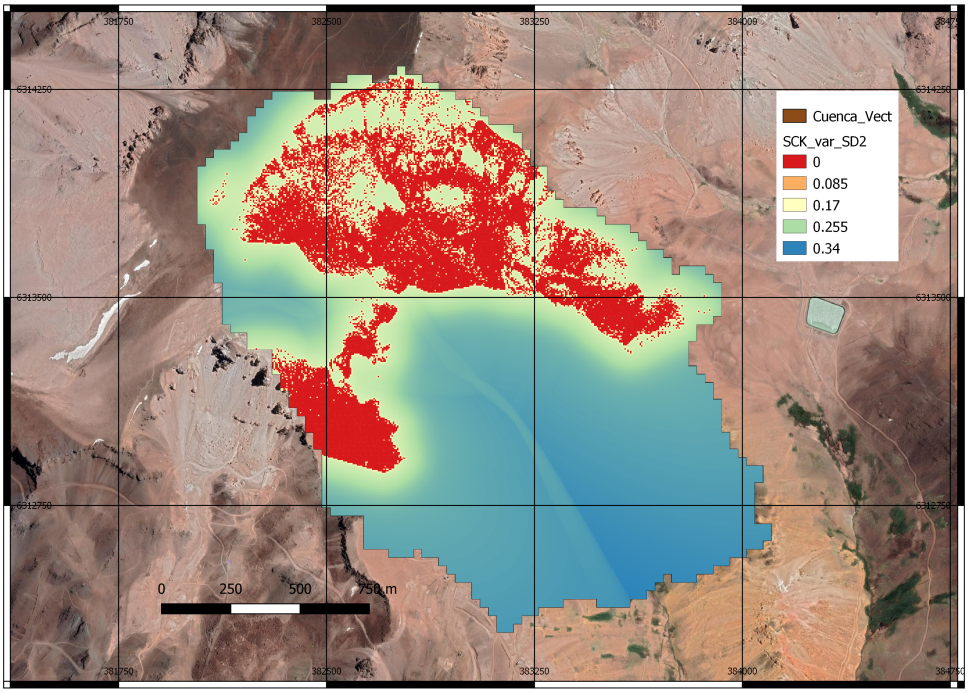


Figura 5.10: Varianza del error de la profundidad de nieve del modelo SCK

En la Figura 5.10 se nota que la varianza del error de la predicción es igual a 0 en los pixeles que se contiene información por lo cual se cumple la condición de exactitud. También se nota que a medida que la predicción se aleja de las mediciones aumenta la varianza del error.

Capítulo 6

Discusión y Análisis de resultados

La tecnología LiDAR ha sido utilizada ampliamente en estudios de hidrología de nieve, diferentes autores han utilizado esta tecnología sobre un avión (ALS por sus siglas en ingles) [20] [15] [28] [7], siendo esta metodología una de las más útiles a la hora de hacer modelos debido a que el producto de la medición entrega un mapa continuo de la profundidad de nieve en toda la cuenca. La desventaja de esta metodología es que es excesivamente costosa.

Para el presente trabajo de título se utilizan mediciones de LIDAR terreste (TLS), la desventaja de este método es que entrega información solo del horizonte de visión del instrumento en terreno. La Figura 3.3 muestra la medición para la fecha sin nieve, en la figura se puede apreciar que las mediciones presentan discontinuidades debido a la posición en la cual se realizó la medición.

A partir del DEM de la cuenca se determinan todas las variables independientes ocupadas en los modelos, un DEM incompleto puede llevar a errores en la generación de las variables independientes. Revuelto et al 2014 [26] utilizan TLS para realizar modelos de árboles de regresión para distintas fechas en los Pirineos, en España. En el estudio, los autores rellenan el DEM utilizando otro DEM obtenido del Instituto Geográfico Nacional de España de manera de no tener que lidiar con las discontinuidades presentes en el DEM.

Para los modelos de ML se utiliza un set de datos correspondientes a píxeles en donde se tiene toda la información, tanto de las variables independientes como de los objetivos. En el proceso de generar los datos se nota que hay partes de la cuenca que no son consideradas debido al muestreo de los datos. Contrario a lo esperado [26] [31], el parámetro Sx muestra poca importancia en los modelos (Tabla 5.2, 5.4, 5.6). La Figura 3.9 muestra el histogram de Sx, en la Figura, se nota que existen valores negativos de la variable, sin embargo, al filtrar los datos para los modelos de ML se nota que ésta posee solo dos valores negativos lo cual explica porque la variable no toma importancia.

Frente al problema del parámetro Sx, es recomendable utilizar un escaneo de la nieve más completo, para así poder abarcar todo el rango de valores que recorre la variable independiente. Las mediciones de profundidad de nieve (Figura 3.15, 3.17, 3.19) poseen discontinuidades y no logran cubrir varias partes de la cuenca, por lo cual una mejora en estas mediciones (Ej:

Realizar dos escaneos de diferentes posiciones en el mismo día) puede ayudar en la mejora de los modelos.

En la literatura, se encuentran diferentes estudios que abordan el problema de modelar la distribución espacial de la profundidad de nieve a una escala interanual [2] [22]. Para el presente estudio se tiene la posibilidad de analizar la influencia de los distintos parámetros para las 3 fechas modeladas, esto entrega información de qué procesos o fenómenos explican de mejor manera la distribución del manto nival a lo largo de la temporada.

La Tabla 5.4 muestra que para la primera fecha la curvatura del terreno es la variable independiente de mayor importancia (30 %), mientras que para las otras dos fechas el TPI es el parámetro más influyente (42 % en cada fecha). TPI analiza de cierta manera la curvatura de el terreno, pero a una escala más grande [26], de esta manera se demuestra que para caracterizar la profundidad de nieve del manto es necesario considerar diferentes escalas. Esto también se aprecia en la Tabla 3.4 donde se analizan diferentes escalas del TPI y se nota que cada escala guarda diferentes correlaciones con la profundidad de nieve.

La distribución de la nieve ha sido modelada por métodos estadísticos como regresiones lineales [17] y aun más ampliamente mediante árboles de regresión [14] [4]. Sin embargo, escasa información existe en el uso de metodologías de ML para modelar la profundidad de nieve [23] [27]. Los resultados expuestos muestran la potencialidad de los modelos de ML para modelar la distribución de la nieve en base a parámetros topográficos, además se rompe el paradigma de la caja negra de los modelos de ML que solo entregan resultados, pues al ser tratado como modelo estadístico se logra entender la importancia relativa de las variables independientes para los modelos. Finalmente, el desempeño de los modelos dan resultados satisfactorios pues entre todos los modelos de ML se logra explicar 37-48 % de la variabilidad de la profundidad de la nieve, rango que se encuentra dentro de lo que es el estado del arte del problema (Tabla 6.1).

Una ventaja que posee el modelo de co-kriging frente al modelo de ML es que en el modelo geoestadístico se logran utilizar todos los datos LiDAR disponibles. Para ML se deben filtrar los datos y luego dividirlos en conjunto de entrenamiento y de prueba, lo que hace que no se utilicen todos los datos de los que se dispone. En este sentido el co-kriging tiene la capacidad de utilizar todos los datos disponibles para calcular variogramas experimentales y resolver las ecuaciones de co-kriging en la vecindad.

La cuenca experimental Piuquenes posee una red de estaciones instaladas que miden tanto la profundidad de nieve, dirección y velocidad del viento, humedad relativa, entre otros parámetros meteorológicos. Para el presente trabajo se utilizan las estaciones solo para caracterizar la dirección predominante del viento. El modelo geoestadístico posee la potencialidad de incorporar las mediciones de las estaciones meteorológicas dentro del modelo mediante la agregación de una nueva variable externa. En este sentido el modelo de co-kriging es mucho más flexible pues si se conocen los variogramas directos y cruzados, se puede incorporar información fácilmente.

Una posible mejora al modelo de co-kriging propuesto seria mejorar el algoritmo de búsqueda de datos en la vecindad del co-kriging. El modelo propuesto busca la información más cercana para resolver la optimización de los pesos, en este sentido ésta puede no ser siempre

Tabla 6.1: Resumen de modelos de distribución de profundidad de nieve

Autor	Tipo de Modelo	r2	Observación
Revuelto et al 2014 [26]	MLR	0.25 - 0.65	TLS
	BRT	0.39	
Balk et al 2000 [3]	BRT	0.54 - 0.65	Campaña de Medición
Molotoch et al 2005 [25]	BRT	0.37	Campaña de Medición
Erxelben et al 2002 [14]	BRT	0.25	Campaña de Medición
Elder et al 1998 [10]	BRT	0.40	Campaña de Medición
Grunewald et al 2013 [17]	MLR	0.27 - 0.90	ALS
Windstral et al 2002 [31]	BRT	0.50	Campaña de Medición
Marofi et al 2011 [23]	ANN	0.75	Campaña de Medición

la mejor opción pues las co-variables pueden ser insuficientes para explicar la variable de interés. Una opción es hacer una búsqueda sectorizada en la cual se da prioridad buscar la variable de interés hasta cierto alcance y luego buscar las co-variables más cercanas. Esta mejora también presenta utilidad debido a que se podría incorporar la información de las estaciones meteorológicas que se tienen dentro de la cuenca.

Los modelos geoestadísticos se utilizan principalmente para la interpolación de variables. La figura 5.10 muestra la varianza esperada del error de co-kriging, en la imagen se nota que la varianza aumenta rápidamente a medida que se aleja de la medición. Esto muestra que la capacidad de extrapolación del modelo de co-kriging es limitada y debe usarse más bien para la interpolación.

Una potencialidad importante del modelo geoestadístico propuesto es la simulación. Los modelos de ML entregan resultados estáticos, esto es, el resultado que entrega el modelo es invariante. Los modelos geoestadísticos poseen la potencialidad de generar muchas simulaciones a partir del mismo set de datos, con lo que se podría obtener diferentes escenarios de la distribución de la profundidad de nieve para una misma fecha. Este trabajo queda abierto pues se escapa de los alcances de la memoria.

Diferentes autores han investigado el efecto de distribuir los residuales de los modelos estadísticos con modelos de kriging y co-kriging [3] [32]. Se investigó esta opción para los modelos de ML (no se muestra el resultado), logrando una mejora del r^2 del 2 %, esto debido a que principalmente, los errores de los modelos no se encontraban fuertemente correlacionados espacialmente. Una mejora del 2 % se considera un buen resultado pues se encuentra dentro de los rangos de mejora esperables. Molotoch et al (2005) [25] logran mejorar entre un 1 y 8 % sus modelos mediante esta técnica mientras que Erxelben et al 2002 [14] solo logran un 3 % de mejora.

Finalmente como última discusión, se destaca que los modelos ajustados responden mas bien a una calibración local y debería desarrollarse una metodología si se quiere aplicar en otro lugar. Grunewald et al 2013 [17] analizan diferentes cuencas alrededor del mundo e intentan ajustar un modelo global que se ajuste a todas. En el estudio, los autores muestran que existen claras diferencias entre los modelos estadísticos hechos para cada cuenca y que un modelo global que incorpore todas las cuencas no represente de buena manera la distribución espacial de la nieve (23-30 % de la variabilidad). Por otra parte, en el estudio, también se llega a la conclusión que los modelos locales ajustados para un año, pueden ser aplicados de buena manera para otros años para las fechas de máxima acumulación.

Conclusión

A lo largo del trabajo se muestran diferentes modelos estadísticos que utilizan de base tecnología LiDAR TLS, se destacan además las utilidades, ventajas y desventajas de esta tecnología. Se concluye que esta tecnología posee un enorme potencial para recopilar información del terreno de manera rápida y precisa. Sin embargo, se destaca que el TLS solo logra obtener información en parte de la cuenca y deja vacíos que condicionan tanto las variables independientes como el desempeño de los modelos.

La distribución de la profundidad de nieve depende fuertemente de la topografía del lugar. Con los modelos propuestos se logra cuantificar la influencia de cada parámetro encontrando. Para la primera fecha, la Curvatura posee la mayor influencia en los modelos y para la segunda y tercera fecha de medición el TPI con un alcance de 100 metros muestra ser la variable de mayor peso en los modelos.

Se compara el desempeño de ML con los modelos convencionales de distribución de la profundidad de la nieve (Tabla 6.1) y se llega a la conclusión de que los modelos propuestos se encuentra dentro de los rangos del estado del arte del problema y representan de manera satisfactoria la variabilidad espacial de le profundidad de nieve en la cuenca Piuquenes.

Se utiliza la metodología "Leave One Out" para analizar el desempeño de los modelos de co-kriging. Los modelos de co-kriging muestran un gran desempeño en una vecindad cercana a los datos LiDAR. La Figura 5.10 muestra como aumenta la varianza del error esperada a medida que la predicción se aleja de la zona de datos LiDAR, mostrando que la principal ventaja de estos modelos consiste en la interpolación y no en la extrapolación.

El modelo geoestadístico mostrado en el trabajo muestra una gran flexibilidad en el sentido de que logra utilizar toda la información LiDAR de la que se dispone y muestra el potencial de incluir en el modelo información externa. No así los modelos de ML, para la metodología mostrada hay información que no se utiliza y depende fuertemente de los datos que se tengan disponibles.

Bibliografía

[1] André Altmann, Laura Toloşi, Oliver Sander, and Thomas Lengauer. Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10):1340–1347, 2010.

[2] SP Anderton, SM White, and B Alvera. Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes*, 18(3):435–453, 2004.

[3] Benjamin Balk and Kelly Elder. Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. *Water Resources Research*, 36(1):13–26, 2000.

[4] Ignacio Moreno Baños, Antoni Ruiz García, Jordi Marturià i Alavedra, Pere Oller, I Figueras, Jordi Piña Iglesias, Carles Garcia Sellés, Pere Martínez I Figueras, and Julià Talaya López. Snowpack depth modelling and water availability from lidar measurements in eastern pyrenees.

[5] Mariana Belgiu and Lucian Drăguţ. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:24–31, 2016.

[6] C Crisci, B Ghattas, and Ghattas Perera. A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*, 240:113–122, 2012.

[7] Jeffrey S Deems, Steven R Fassnacht, and Kelly J Elder. Fractal distribution of snow depth from lidar data. *Journal of Hydrometeorology*, 7(2):285–297, 2006.

[8] Jeffrey S Deems, Thomas H Painter, David C Finnegan, et al. Lidar measurement of snow depth: a review. *J. Glaciol*, 59(215):467–479, 2013.

[9] Olivier Dubrule. Cross validation of kriging in a unique neighborhood. *Journal of the International Association for Mathematical Geology*, 15(6):687–699, 1983.

[10] Kelly Elder, Walter Rosenthal, and Robert E Davis. Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes*, 12(10-11):1793–1808, 1998.

[11] Xavier Emery. Iterative algorithms for fitting a linear model of coregionalization. *Com-*

puters & Geosciences, 36(9):1150–1160, 2010.

[12] Xavier Emery. Cokriging random fields with means related by known linear combinations. *Computers & geosciences*, 38(1):136–144, 2012.

[13] Tyler A Erickson, Mark W Williams, and Adam Winstral. Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, colorado, united states. *Water Resources Research*, 41(4), 2005.

[14] Jennifer Erxleben, Kelly Elder, and Robert Davis. Comparison of spatial interpolation methods for estimating snow distribution in the colorado rocky mountains. *Hydrological Processes*, 16(18):3627–3649, 2002.

[15] SR Fassnacht and JS Deems. Measurement sampling and scaling for deep montane snow depth data. *Hydrological Processes: An International Journal*, 20(4):829–838, 2006.

[16] Jerome H Friedman. Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4):367–378, 2002.

[17] T Grunewald, J Stotter, John W Pomeroy, Ruzica Dadic, I Moreno Banos, J Marturià, Maximilian Sproß, Christopher Hopkinson, Paolo Burlando, and M Lehnig. Statistical modelling of the snow depth distribution in open alpine terrain. 2013.

[18] Tobias Jonas, Christoph Marty, and Jan Magnusson. Estimating the snow water equivalent from snow depth measurements in the swiss alps. *Journal of Hydrology*, 378(1-2):161–167, 2009.

[19] Mikhail Kanevski, Vadim Timonin, and Alexi Pozdnukhov. *Machine learning for spatial environmental data: theory, applications, and software*. EPFL press, 2009.

[20] PB Kirchner, RC Bales, NP Molotch, J Flanagan, and Q Guo. Lidar measurement of seasonal snow accumulation along an elevation gradient in the southern sierra nevada, california. *Hydrology and Earth System Sciences*, 18(10):4261–4275, 2014.

[21] John B Lindsay. Whitebox gat: A case study in geomorphometric analysis. *Computers & Geosciences*, 95:75–84, 2016.

[22] Juan I López-Moreno, SR Fassnacht, JT Heath, KN Musselman, Jesús Revuelto, Jérôme Latron, Enrique Morán-Tejeda, and Tobias Jonas. Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. *Advances in Water Resources*, 55:40–52, 2013.

[23] Safar Marofi, Hossein Tabari, and Hamid Zare Abyaneh. Predicting spatial distribution of snow water equivalent using multivariate non-linear regression and computational intelligence methods. *Water resources management*, 25(5):1417–1435, 2011.

[24] Vivian Meløysund, Bernt Leira, Karl V Høiseth, and Kim Robert Lisø. Predicting snow density using meteorological data. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, 14(4):413–423, 2007.

[25] NP Molotch, MT Colee, RC Bales, and J Dozier. Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrological Processes: An International Journal*, 19(7):1459–1479, 2005.

[26] Jesús Revuelto, Juan I López-Moreno, César Azorin-Molina, and Sergio M Vicente-Serrano. Topographic control of snowpack distribution in a small catchment in the central spanish pyrenees: intra-and inter-annual persistence. *The Cryosphere*, 8(5):1989–2006, 2014.

[27] Hossein Tabari, S Marofi, H Zare Abyaneh, and MR Sharifi. Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in samsami basin of iran. *Neural Computing and Applications*, 19(4):625–635, 2010.

[28] Wade T Tinkham, Alistair MS Smith, Hans-Peter Marshall, Timothy E Link, Michael J Falkowski, and Adam H Winstral. Quantifying spatial distribution of snow depth errors from lidar using random forest. *Remote sensing of environment*, 141:105–115, 2014.

[29] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, and Alexis Fouilloy. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582, 2017.

[30] Karl Wetlaufer, Jordy Hendrikx, and Lucy Marshall. Spatial heterogeneity of snow density and its influence on snow water equivalence estimates in a large mountainous basin. *Hydrology*, 3(1):3, 2016.

[31] Adam Winstral, Kelly Elder, and Robert E Davis. Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of hydrometeorology*, 3(5):524–538, 2002.

[32] Zeshi Zheng, Noah P Molotch, Carlos A Oroza, Martha H Conklin, and Roger C Bales. Spatial snow water equivalent estimation for mountainous areas using wireless-sensor networks and remote-sensing products. *Remote sensing of environment*, 215:44–56, 2018.