

**SOCIEDAD CHILENA DE INGENIERIA HIDRÁULICA.
XXV CONGRESO CHILENO DE INGENIERIA HIDRÁULICA.**

**COMPARACIÓN DE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO VERSUS UN
MODELO DE PROCESOS EN LA ESTIMACIÓN DE CAUDALES. APLICACIÓN A
UNA CUENCA DE ALTA MONTAÑA.**

BENJAMIN GONZALEZ CALVO.¹

LUIS DE LA FUENTE C..²

LINA CASTRO.³

RESUMEN

Las cuencas de alta montaña son fundamentales para la gestión hídrica en Chile, pero su complejidad climática y comportamiento altamente variable suponen desafíos para los modelos hidrológicos tradicionales. Frente a ello, los modelos de Machine Learning (ML) permiten capturar relaciones no lineales y descifrar patrones de la cuenca sin definir ecuaciones. Este estudio evalúa el desempeño de tres modelos de ML, Random Forest, LSTM y CNN, frente al modelo hidrológico USM, en la predicción de caudales aplicados a cuenca de alta montaña. Se utilizaron variables meteorológicas y radiación extraterrestre, organizadas en ventanas rezagadas y agregadas para representar la memoria hidrológica. El modelo LSTM alcanzó el mejor desempeño global con un KGE de 0.80, seguido por CNN el cual obtuvo un valor 0.72 y RF con un valor de 0.72, superando en todos los casos al modelo hidrológico USM. A nivel interanual, se identificaron diferencias en la capacidad de los modelos para representar eventos extremos, siendo el LSTM el que mostró mayor consistencia. Estos resultados destacan el potencial de los modelos de ML como herramientas predictivas robustas para escenarios hidrológicos complejos, aportando evidencia para futuras aplicaciones en la planificación y gestión hídrica.

¹ Estudiante, Ingeniería Civil, Universidad Federico Santa María - email: benjamin.gonzalezc@sansano.usm.cl

² Profesor Guía, Departamento de Obras Civiles, Universidad de Concepción - email: ladelafuenteco@utep.edu

³ Profesora Co-guía, Departamento de Obras Civiles, Universidad Federico Santa María - email:
Lina.Castro@usm.cl

1. Introducción

Las cuencas de alta montaña cumplen un rol clave en el abastecimiento de agua potable y de riego en poblaciones aguas abajo, actúan como reservorios naturales que regulan el régimen hídrico anual (Blöschl et al., 2019). La predicción de caudales en cuencas de alta montaña representa un desafío significativo para la modelación hidrológica debido a la complejidad de los procesos involucrados. En estas cuencas, la precipitación, el deshielo nival y glaciario son determinantes clave en la escorrentía, lo que las convierte en sistemas altamente variables y sensibles a cambios climáticos (Sanllehi, 2023).

Tradicionalmente, los modelos hidrológicos de procesos físicos han sido la herramienta principal para la estimación de caudales, basándose en ecuaciones que representan procesos como la escorrentía, infiltración y la evapotranspiración (Blöschl et al., 2019). Sin embargo, estos modelos requieren complejas calibraciones y sus ecuaciones simplifican considerablemente la realidad hidrológica (Singh et al., 2023).

En los últimos años, las técnicas de Machine Learning (ML) han ganado reconocimiento en la hidrología debido a su capacidad de manejar grandes volúmenes de datos y descifrar patrones complejos sin depender de ecuaciones predefinidas (Rozos, 2023). Modelos como Long Short Term Memory (LSTM) han demostrado superar a modelos hidrológicos tradicionales en predicción de caudales (Kratzert et al., 2019), mientras que otros estudios han mostrado que incluso técnicas más simplificadas, como Random Forest (RF) y Convolutional Neural Networks (CNN), presentan mejores desempeños que los modelos tradicionales (Yang et al., 2019; Singh et al., 2023).

Este estudio evalúa si modelos de ML, como Random Forest, LSTM y CNN, pueden alcanzar un desempeño comparable o superior al modelo hidrológico USM en la predicción de caudales medios diarios, el modelo hidrológico USM es un modelo altamente calibrado en el estudio de Sanllehi (2023). Además, se analizan las implicancias de sus errores y aciertos en años secos, húmedos y extremos, para contribuir a una comprensión más profunda de su aplicabilidad en sistemas hidrológicos complejos (Anilkumar et al., 2023).

2. Metodología

2.1. Zona de estudio

La cuenca del río Juncal, ubicada en la Región de Valparaíso (Chile), fue seleccionada como caso de estudio debido a su importancia en el abastecimiento hídrico del río Aconcagua y a su carácter representativo de las cuencas de alta montaña de la zona central del país. Con una superficie de 343 km² y un 7,5% de su territorio cubierto por glaciares, presenta un régimen nival marcado, con máximos caudales en primavera y verano, asociados al derretimiento de nieve y hielo.

2.2. Modelo hidrológico USM

Propuesto por Stowhas (1975), en el esfuerzo por mejorar la representación del derretimiento nival y glaciario en las cuencas de alta montaña de Chile central, la investigación de Sanllehi Placencia (2023) propuso una reestructuración del modelo hidrológico USM. Este enfoque permitió una estimación más precisa de los aportes hídricos provenientes de la nieve, los glaciares y la

precipitación líquida. Para ello, se integraron tres modelos, cada uno con un rol específico en la simulación del ciclo hidrológico de las cuencas que analizó.

2.2.1 Modelo de Balance de Energía de Utah (UEB):

Para mejorar las limitaciones del modelo USM en su capacidad de representar los procesos energéticos asociados al manto nival, incorporaron el Modelo de Balance de Energía de Utah, un enfoque basado en las ecuaciones de balance energético y de masa. Este modelo permite estimar con mayor precisión la evolución del manto de nieve al considerar factores determinantes como la radiación solar, la temperatura del aire y el contenido energético de la nieve.

2.2.2 Modelo de Grado-Índice Mejorado (ETI):

Si bien el derretimiento nival constituye una fuente predominante de escorrentía, el retroceso glaciar y su contribución al caudal de los ríos se han vuelto cada vez más relevantes en el contexto de cambio climático. Para abordar este componente, Sanllehi integró el Modelo de Grado e Índice Mejorado (ETI), que introduce una combinación del enfoque clásico de grado-día con la radiación de onda corta.

El modelo ETI incorpora la radiación solar como un impulsor clave del proceso de fusión. Este cambio permitió mejorar la estimación de los aportes glaciales al caudal y evaluar su importancia relativa dentro del balance hídrico de las cuencas analizadas.

2.2.3 Modelo de Escorrentía de Migueles:

Finalmente, la estimación de los aportes hídricos no estaría completa sin un esquema robusto para simular la escorrentía generada por la combinación de derretimiento nival, glaciar y la precipitación líquida. Para ello, se implementó el Modelo de Escorrentía de Migueles, un modelo conceptual, de características semi distribuido diseñado para representar la transformación del agua disponible en caudal efectivo.

Este modelo permite procesar de manera integrada los diferentes aportes hídricos y simular su transporte a través del sistema hidrológico. Las variables clave consideradas en su formulación incluyen la cantidad de agua derretida, la infiltración del suelo y el escurrimiento superficial y subterráneo.

2.2.4 Calibración Modelo Hidrológico USM

El modelo USM fue previamente calibrado en el estudio de Sanllehi (2023), la cual se tomó como estudio de referencia en este trabajo. Su calibración se realizó por etapas, considerando fuentes de información específicas para cada componente del modelo. Por ejemplo, se utilizaron imágenes satelitales MODIS para el ajuste del módulo nival, tasas de cambio de elevación glaciar derivadas de imágenes ASTER para el módulo glaciar, y registros de caudal en puntos de aforo para el módulo de escorrentía.

En cuanto a los métodos empleados, se utilizó el algoritmo de evolución compleja SCE-UA para la optimización de los parámetros, maximizando funciones objetivo como la eficiencia de Kling-Gupta (KGE) o el coeficiente de determinación (R^2), según el tipo de variable evaluada. Además, se definieron escenarios de calibración y validación combinando años secos y húmedos, con el fin de representar distintos regímenes hidrológicos. Para el caso del módulo de escorrentía, se probaron cuatro combinaciones distintas de configuración (variando el número de bandas y el rango del

parámetro HSM_{max}), lo que permitió seleccionar el escenario con mejor eficiencia en cada cuenca.

Cabe señalar que en este estudio no se modificó ni recalibró el modelo USM, el cual se mantuvo como línea base para efectos comparativos frente a los modelos de aprendizaje automático.

2.3 Base de datos

Para evaluar de manera equitativa el desempeño de los modelos de ML y el modelo Hidrológico USM, se utilizaron series diarias comprendidas entre los años 2000 a 2019. La base de datos incluye caudales observados y variables meteorológicas claves, descritas a continuación:

- Precipitación diaria (mm): Se utilizaron datos del producto satelital CR2MET (Alvarez-Garretón et al., 2018), con resolución espacial de 0.05° y frecuencia diaria. La precipitación es una de las variables más críticas en el balance hídrico, ya que determina directamente el volumen de agua disponible en la cuenca.
- Temperatura máxima y mínima ($^\circ\text{C}$): Extraída de CR2MET, esta variable es clave en la modelación del derretimiento nival y glaciar, además de influir en la evapotranspiración.
- Velocidad del viento (m/s): Datos con frecuencia horaria, ajustados mediante el Explorador de Energía Eólica, esenciales para la evaporación y el transporte de humedad.
- Radiación solar extraterrestre (W/m^2): Esta variable se incorpora como indicador de energía solar incidente potencial. Se calcula a partir de las coordenadas geográficas y el día juliano siguiendo el procedimiento de Allen et al. (1998) y Iqbal et al. (1983), considerando la declinación solar y la distancia Tierra-Sol. Este cálculo permite construir una serie diaria continua que refleja la disponibilidad energética para evapotranspiración y fusión nival.
- Caudales diarios observados (m^3/s): Se obtienen de la Dirección General de Aguas (DGA), representando los valores reales medidos en la cuenca. Estos datos son clave en el entrenamiento de los modelos de ML, ya que vamos a ingresar estos caudales como la variable a predecir para cada modelo.

2.3.1 Construcción de variables derivadas y partición de datos

Para capturar los procesos hidrológicos que ocurren a diferentes escalas temporales, se construyen dos tipos de variables derivadas. En primer lugar, se generaron variables rezagadas de 1 a 7 días, las cuales permiten representar la memoria corta de la cuenca, es decir, el impacto directo e inmediato de los eventos recientes (por ejemplo, tormentas o eventos que se reflejen aumentos súbitos de temperatura).

En segundo lugar, se calculan variables agregadas en ventanas móviles más amplias, utilizando la información semanal, repartidas en ventanas de 8–15, 16–31, 32–62, 63–123 y 124–224 días. Estas variables de memoria larga capturan tendencias estacionales y efectos acumulativos en el almacenamiento de agua en nieve, glaciares, suelos y acuíferos.

Para garantizar una comparación equitativa entre los modelos de ML y el modelo hidrológico USM, se utilizó el mismo criterio de división de datos propuesto por Sanllehi (2023). La serie de tiempo

completa fue dividida en tres subconjuntos, entrenamiento, validación y prueba final. En particular, los años utilizados para el entrenamiento fueron 2002, 2003, 2004, 2009, 2011, 2012, 2014 y 2017.

Mientras que los años de validación correspondieron a los años 2006 y 2018, finalmente los datos de la prueba final abarcan los años 2000, 2001, 2005, 2007, 2008, 2010, 2013, 2015, 2016 y 2019.

Esta clasificación fue diseñada para incluir tanto años secos como húmedos en cada conjunto, asegurando que en el entrenamiento y evaluación de los modelos consideraran la variabilidad climática de la cuenca.

En base al conjunto de entrenamiento se debe realizar la búsqueda de hiperparámetros para cada modelo, ajustando sus configuraciones internas con el objetivo de mejorar su capacidad de generalización y evitar sobreajuste. Posteriormente, los modelos calibrados se validan y se evalúan sobre conjunto de prueba final, garantizando una comparación robusta y sin sesgos.

2.4 Modelos de ML

2.4.1 Random Forest (RF)

El modelo Random Forest (RF), propuesto por Breiman (2001), es un método de aprendizaje supervisado basado en un ensamble de múltiples árboles de decisión. Su fortaleza radica en que cada árbol se entrena sobre una muestra diferente de los datos y con subconjuntos aleatorios de variables predictoras, lo que reduce la varianza de las predicciones finales y mejora la estabilidad del modelo.

Predicción de un Árbol de Decisión: Cada árbol individual realiza una predicción \hat{y}_i para una entrada x basada en su estructura interna, definida por la función de decisión $f(x)$:

$$\hat{y}_i = f(x) \quad (1)$$

La predicción final de RF se obtiene promediando las predicciones de todos los árboles

$$y = \frac{1}{y_0} \sum_{y_0=1}^{y_0} \hat{y}_i \quad (2)$$

2.4.2 Long Short-Term Memory (LSTM)

La red neuronal Long Short-Term Memory (LSTM), propuesta por Hochreiter & Schmidhuber (1997), es una arquitectura recurrente diseñada para modelar dependencias de largo plazo en series temporales. Su característica principal es la incorporación de unidades de memoria que controlan qué información se almacena, se olvida y se utiliza en cada paso de tiempo. Las unidades de memoria que regulan el flujo de información mediante tres compuertas.

Puerta de entrada: Controla la cantidad de información nueva que se agrega a la celda de memoria.

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + B_i) \quad (3)$$

donde i_t es la activación de la puerta de entrada, W_i son los pesos, h_{t-1} es la salida de la celda anterior, x_t es la entrada actual y B_i es el sesgo.

Puerta de olvido: Decide qué información se debe eliminar de la celda de memoria.

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + B_f) \quad (4)$$

donde f_t es la puerta del olvido para el tiempo t .

Puerta de Salida: Controla la salida de la celda de memoria.

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + B_o) \quad (5)$$

donde o_t es la puerta de salida.

La celda de memoria se actualiza de la siguiente manera:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

donde C_t es el estado de la celda en el tiempo t .

Finalmente, la salida de la LSTM se calcula como:

$$h_t = o_t * \text{Tanh}(C_t) \quad (7)$$

2.4.3 Convolutional Neural Networks (CNN)

Desarrollado por Lecun (1998), corresponden a redes neuronales diseñadas originalmente para el procesamiento de imágenes, pero adaptadas aquí para extraer patrones temporales en las series, mediante un filtro unidimensional (1D-CNN). La CNN aplica filtros de convolución que barren las ventanas de tiempo, detectando tendencias locales y recurrentes.

$$y_{(i,j)} = \sum_{m=1}^k \sum_{n=1}^k x_{(i+m-1,j+n-1)} * w_{(m,n)} + b \quad (8)$$

- $x_{(i+m-1,j+n-1)}$: Representa los valores de la entrada (como una imagen o serie de datos).
- $w_{(m,n)}$: Es el kernel o filtro que se desplaza por la entrada.
- k : Es el tamaño del filtro.
- b : Es el sesgo que permite ajustar la salida.
- $y_{(i,j)}$: Es el valor del mapa de características en la posición (i,j) .

Para introducir no linealidad en el modelo, las CNN emplean funciones como “Rectified Linear Unit” (ReLU), conocida también como función de activación. Las ReLU ayudan a evitar problemas como el desvanecimiento del gradiente en las redes profundas y permiten que la red aprenda relaciones complejas entre las características extraídas.

$$f(x) = \text{Max}(0, x) \quad (9)$$

Después de esta fase, se emplea una técnica llamada max-pooling, que selecciona los valores más representativos de cada ventana de tiempo, reduciendo la cantidad de información sin perder los detalles más importantes. Finalmente, la red neuronal combina todos estos datos procesados para generar una predicción del caudal esperado en función de los patrones que ha aprendido a lo largo del entrenamiento.

2.5 Entrenamiento de Modelos de Machine Learning

Cada uno de los modelos de Machine Learning fue entrenado utilizando el conjunto de datos previamente dividido en tres periodos, entrenamiento, validación interna y test final. El proceso de entrenamiento se centró en ajustar los hiperparámetros mediante una búsqueda tipo Grid search, empleando el coeficiente Kling-Gupta Efficiency (KGE) como métrica principal de evaluación dentro del conjunto de validación interna. En el caso de Random Forest, se exploraron combinaciones de profundidad máxima del árbol (max_depth), número de árboles (n_estimators) y fracción de variables seleccionadas en cada división (max_features). Para el modelo LSTM, la búsqueda consideró el número de capas (n_layers), la longitud de la ventana de tiempo (seq_length), la cantidad de unidades ocultas (hidden_size), la tasa de aprendizaje (learning_rate) y el valor de regularización tipo dropout (dropout). En el caso de la red convolucional CNN, los hiperparámetros evaluados incluyeron el tamaño del kernel (kernel_size), número de filtros, tasa de dropout (dropout) y tasa de aprendizaje (learning_rate).

2.6 Evaluación de Desempeño

Para cuantificar el desempeño de los modelos en la predicción de caudales medios diarios, se emplearon tres métricas complementarias que permiten evaluar distintos aspectos del error de simulación. La métrica principal seleccionada fue el Kling-Gupta Efficiency (KGE) (Gupta et al., 2009), ampliamente utilizada en hidrología por su capacidad de descomponer el error en componentes interpretables, lo que permite una mejor comprensión del ajuste logrado por los modelos. Como apoyo, también se utilizaron el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE).

2.6.1 Kling-Gupta Efficiency (KGE)

Propuesto por Gupta et al. (2009), busca superar las limitaciones del clásico coeficiente de eficiencia de Nash-Sutcliffe (NSE), particularmente su sensibilidad al sesgo y su tendencia a subestimar la variabilidad del caudal. A diferencia del NSE, el KGE permite descomponer el error en tres componentes, la correlación lineal entre los valores observados y predichos (r), el sesgo medio (β), y la dispersión relativa (α), facilitando así una evaluación más robusta y transparente.

La expresión general de la KGE se define como:

$$KGE = \sqrt{1 - (r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (10)$$

Valores de KGE cercanos a 1 indica una predicción precisa y equilibrada, mientras que valores negativos sugieren una predicción menos fiable.

3 Resultado y discusión

3.1 Resultado de Arquitecturas

Tal como se explicó en el apartado metodológico, cada modelo de ML fue entrenado y validado internamente para identificar la combinación de hiperparámetros que ofreciera el mejor desempeño mediante el método Grid Search, maximizando el valor de KGE. A partir de este proceso, se definieron las arquitecturas finales para cada modelo, las cuales se describen a continuación.

Para el modelo Random Forest, se obtuvo un total de 100 árboles, con una profundidad máxima indefinida, una división mínima de dos muestras.

En el caso del modelo LSTM, la arquitectura final seleccionada consideró dos capas compuestas de 4 neuronas, una secuencia de entrada de 180 días, una tasa de dropout de 0.1 y una learning_rate de 0.001.

Finalmente, CNN fue configurada utilizando el optimizador de Adam, con filtros de tamaño 32 y 64, kernel de 7, una ventana de entrada de 224 días, dropout de 0.1 y la misma tasa de aprendizaje que el modelo LSTM.

3.2 Desempeño General de los modelos

La siguiente tabla resume los resultados globales de desempeño para los cuatro modelos evaluados en el conjunto de la prueba final.

Tabla 1. Resultado de métricas globales para el conjunto de prueba final.

Resultado Métricas	RF	LSTM	CNN	Modelo hidrológico USM
KGE	0.72	0.8	0.72	0.54
RSME	2.37	2.33	2.95	4.23
MAE	1.39	1.59	1.89	2.73

En base a los resultados se observa que LSTM alcanza el mejor desempeño general, con el mayor KGE (0.80) y el menor RMSE, destacándose por su capacidad de reproducir los valores extremos sin comprometer la estabilidad global del modelo. Por otro lado, Random Forest obtiene el menor MAE, lo que indica una excelente precisión media, aunque con menor habilidad para representar variaciones abruptas en los caudales. El modelo CNN logra un KGE comparable al de Random Forest, pero con mayores errores absolutos, lo que sugiere una mayor inestabilidad. Finalmente, el modelo hidrológico USM presenta el desempeño más bajo en todas las métricas, en línea con su menor capacidad para adaptarse a la complejidad del régimen de caudales observado.

Una de las preguntas centrales de esta investigación fue si los modelos de ML pueden igualar o superar el desempeño de un modelo hidrológico conceptual calibrado tradicionalmente. Los resultados confirman que todos los modelos ML obtienen mejores métricas que el modelo USM, tanto en correlación como en reducción de errores. Esta diferencia es especialmente clara en la

métrica KGE y el RMSE, donde se evidencia la dificultad del modelo USM para capturar tanto la sincronía como la magnitud de los caudales extremos.

3.3 Desempeño por año de los modelos

Además del análisis mediante métricas globales, la inspección visual de los hidrogramas anuales permite identificar fortalezas y debilidades específicas de cada modelo. Se utilizaron hidrogramas comparativos en distintos años representativos de la prueba final, los cuales muestran la evolución temporal del caudal observado y simulado por cada modelo.

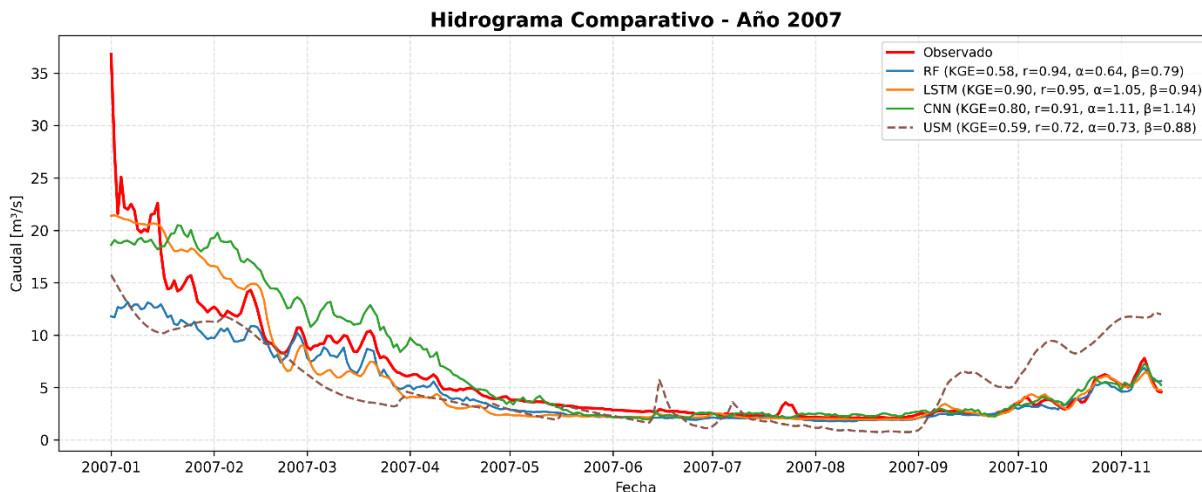


Figura 1. Hidrograma comparativo año 2007 para periodo de prueba final.

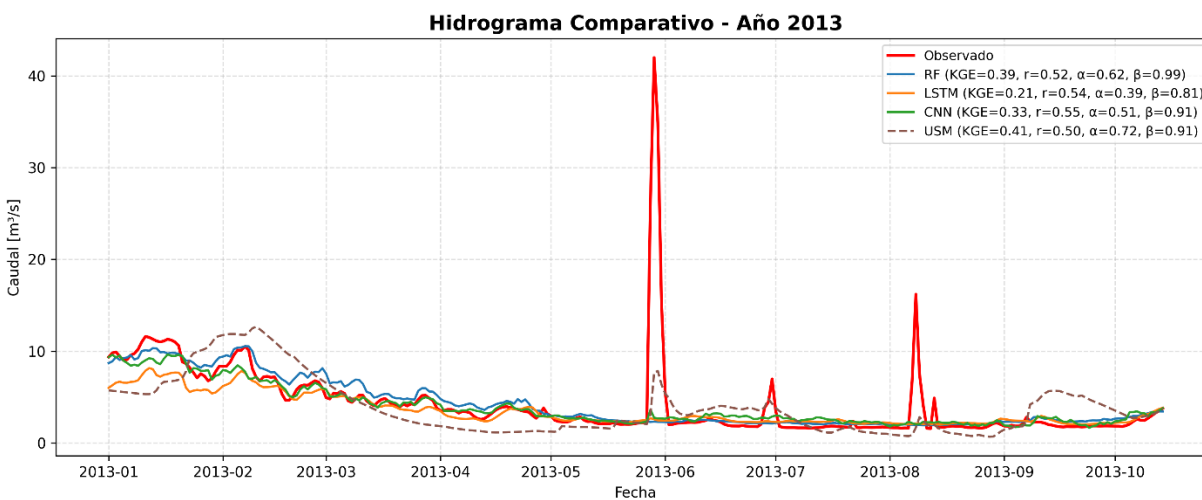


Figura 2. Hidrograma comparativo año 2008 para periodo de prueba final.

Tabla 2. Resultado del KGE por año para el conjunto de prueba final.

KGE por Año	RF	LSTM	CNN	Modelo hidrológico USM
2001	0.64	0.77	0.77	0.53

2005	0.64	0.53	0.81	0.26
2007	0.58	0.9	0.8	0.59
2008	0.9	0.61	-0.65	-0.32
2010	0.78	0.72	0.45	0.48
2013	0.39	0.21	0.33	0.41
2015	0.4	0.54	0.21	0.64
2016	0.7	0.77	0.72	0.35
2019	0.47	0.59	0.06	0.52

De manera general, el modelo LSTM logra reproducir la forma del hidrograma anual con alta fidelidad, capturando tanto los caudales base como las crecidas estacionales, aunque en algunos años sobreestima significativamente las crecidas de caudal. El modelo RF presenta una respuesta más suavizada, reproduciendo adecuadamente la tendencia media, pero con menor precisión en eventos extremos que LSTM. Por su parte, el CNN muestra comportamientos más erráticos, en algunos años su respuesta es sobresaliente (2005 o 2007), pero en otros presenta fluctuaciones artificiales o sobreestimaciones marcadas (2008), lo que evidencia problemas de generalización en el conjunto de datos de entrenamiento y posibles casos de overfitting.

El modelo USM, pese a ser generalmente superado por las técnicas de ML, logró buenos resultados en años específicos (2007 y 2015), e incluso obtuvo el mejor componente β en varios años, lo que indica un adecuado balance en el volumen total de agua anual.

3.4 Discusión

Los resultados obtenidos en este estudio permiten reflexionar sobre el rol actual y potencial de los modelos de ML en contextos de alta montaña. Si bien los modelos ML especialmente LSTM superaron al USM en métricas globales como KGE y RMSE, esta diferencia no fue constante ni absoluta. En algunos años como 2007 o 2015, el modelo USM mostró componentes β cercanos a 1, lo cual indica un buen balance de volumen e incluso, en ciertos casos, un desempeño comparable al de modelos de ML. Esto refuerza que los modelos conceptuales aún tienen un valor operativo, especialmente cuando se dispone de pocos datos o se busca un modelo con estructura física interpretativa.

El valor de esta comparación no solo busca predecir mejor, sino entender qué tipo de información representa cada modelo, qué errores comete, y en qué contextos lo hacen. Por ejemplo, el LSTM mostró una alta capacidad para aprender patrones interanuales y reproducir tanto caudales base como crecidas estacionales. Pero al mismo tiempo, sus errores en eventos extremos revelan que la memoria por sí sola no es suficiente si el evento no está suficientemente representado en el conjunto de datos de entrenamiento.

Por otro lado, Random Forest, sin contar con estructuras secuenciales, entrega un buen ajuste promedio, destacando por su precisión en años estables. CNN logra capturar picos abruptos, pero tiende a ser muy sensible y sobrerreaccionar en ausencia de patrones claros. Estas diferencias no son triviales, ya que muestran que cada modelo aporta al régimen hidrológico una visión distinta,

lo cual es útil para aplicaciones distintas, como por ejemplo la gestión de embalses, alertas tempranas, monitoreo base, etc.

Desde una perspectiva más amplia, si los resultados hubieran mostrado que el USM era consistentemente mejor, probablemente este trabajo habría servido para reafirmar la vigencia de los enfoques físicos, pero al constatar que los modelos de ML pueden replicar o superar su desempeño, se refuerza la necesidad de explorar nuevas estrategias de modelación, especialmente en un contexto de cambio climático e incertidumbre creciente. No obstante, estos resultados corresponden a una sola cuenca de alta montaña, por lo que sería necesario realizar estudios adicionales en otras cuencas y contextos hidrológicos para validar y generalizar las conclusiones obtenidas.

Desde una perspectiva más general, estos hallazgos no buscan invalidar el uso de modelos físicos ni proponer que deban ser reemplazados. Más bien, sugieren que los modelos de ML pueden ser considerados como una herramienta sólida, capaces de encontrar una muy buena respuesta a los caudales de la cuenca de estudio, sin imponer ecuaciones matemáticas y principios físicos, sino que los modelos aprendieron de los datos. Esta complementariedad ofrece oportunidades para desarrollar estrategias de modelación más adaptativas, sin necesidad de abandonar completamente los enfoques clásicos.

4 Conclusión

A partir del análisis de métricas globales y de los hidrogramas anuales comparativos, se concluye que los modelos de aprendizaje automático evaluados presentan un desempeño competitivo frente al modelo hidrológico USM. En términos de la métrica KGE, todos los modelos de ML superan consistentemente al modelo físico en la mayoría de los años, lo que evidencia una mayor capacidad para replicar la señal hidrológica observada y ajustarse muy bien a los valores observados.

Estos resultados indican que, bajo ciertas condiciones de entrenamiento y una adecuada selección de atributos, los modelos de aprendizaje automático no solo pueden igualar, sino también superar el rendimiento de modelos hidrológicos estructurados. Esta ventaja es particularmente relevante en cuencas con alta variabilidad climática o presencia de procesos no lineales difíciles de representar utilizando enfoques físicos clásicos. La capacidad de los modelos ML para incorporar múltiples fuentes de información climática y generar representaciones flexibles de la dinámica del caudal, los posiciona como una alternativa robusta para aplicaciones operativas o en escenarios de cambio climático.

Sin embargo, estos hallazgos no implican necesariamente que los modelos físicos deban ser reemplazados, sino que los modelos de ML constituyen una herramienta complementaria con gran potencial, especialmente en contextos donde la calibración tradicional tiende a ser costosa o limitada. La fortaleza de estas técnicas radica en su capacidad para poder capturar patrones complejos del sistema hidrológico de la cuenca y adaptarse a diferentes condiciones climáticas sin requerir un conocimiento profundo de los procesos físicos locales de la zona. En ese sentido, este estudio demuestra que, en ciertos escenarios, es posible alcanzar desempeños competitivos frente a modelos físicos estructurados, lo que abre oportunidades para integrar enfoques alternativos en futuras investigaciones hidrológicas y aplicaciones operativas. El mensaje final es que los modelos de ML desafían y complementa los enfoques clásicos existentes, ofreciendo nuevas rutas para abordar la predicción de caudales en cuencas complejas.

AGRADECIMIENTOS

Esta memoria se realizó en el marco de dos proyectos de investigación: Proyecto FONDECYT Iniciación en investigación No.11220482 y el proyecto USM Interno Multidisciplinario PI_M_24_03.

REFERENCIAS

- Anilkumar, R., Bharti, R., Chutia, D., & Aggarwal, S. P. (2023). Modelling point mass balance for the glaciers of the Central European Alps using machine learning techniques. *Cryosphere*, 17(7), 2811–2828. <https://doi.org/10.5194/tc-17-2811-2023>
- Duan, S., Ullrich, P., & Shu, L. (2020). Using Convolutional Neural Networks for Streamflow Projection in California. *Frontiers in Water*, 2. <https://doi.org/10.3389/frwa.2020.00028>
- Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2021). Rainfall-runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, 25(4), 2045–2062. <https://doi.org/10.5194/hess-25-2045-2021>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine-Learning Applied to Large-Sample Datasets. <http://arxiv.org/abs/1907.08456>
- Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>
- Rozos, E. (2023). Assessing Hydrological Simulations with Machine Learning and Statistical Models. *Hydrology*, 10(2). <https://doi.org/10.3390/hydrology10020049>
- Sanllehi. (2023). Integración y evaluación de modelo de derretimiento y acumulación de nieve y hielo glaciar a modelo hidrológico usm.
- Shortridge, J. E., Guikema, S. D., & Zaitchik, B. F. (2016). Machine learning methods for empirical streamflow simulation: A comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrology and Earth System Sciences*, 20(7), 2611–2628. <https://doi.org/10.5194/hess-20-2611-2016>
- Singh, D., Vardhan, M., Sahu, R., Chatterjee, D., Chauhan, P., & Liu, S. (2023). Machine-learning- and deep-learning-based streamflow prediction in a hilly catchment for future scenarios using CMIP6 GCM data. *Hydrology and Earth System Sciences*, 27(5), 1047–1075. <https://doi.org/10.5194/hess-27-1047-2023>
- Yang, S., Yang, D., Chen, J., Santisirisomboon, J., Lu, W., & Zhao, B. (2020). A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *Journal of Hydrology*, 590, 125206. <https://doi.org/10.1016/J.JHYDROL.2020.125206>