

Modelos Híbridos Hidrológicos basados en LSTM y simuladores físicos

Hybrid Hydrological Modeling with LSTM and physical simulators

N. Tacoronte (1), M. García-Valdecasas Ojeda (1, 2), Y. Castro-Díez (1,2), M.J. Esteban-Parra (1,2), S. Gámiz-Fortis (1,2)

(1) Depto. Física Aplicada, Universidad de Granada. (2) Instituto Interuniversitario de Investigación del Sistema Tierra en Andalucía (IISTA-CEAMA), Granada

RESUMEN

Los modelos hidrológicos físicos proporcionan datos interpretables y de consistencia física, pero pueden degradarse cuando existen incertidumbres en parámetros, sesgos en forzamientos meteorológicos o simplificaciones estructurales del modelo (omisiones o parametrizaciones simplificadas de procesos como almacenamiento subterráneo, nieve o evapotranspiración). En paralelo, las redes neuronales recurrentes tipo “Long Short-Term Memory” (LSTM) han mostrado capacidad para aprender relaciones no lineales y dependencias temporales en series hidrometeorológicas. En este trabajo se evalúa el potencial de combinar ambos enfoques con un esquema híbrido, analizando la mejora del rendimiento predictivo, y el porqué en algunos casos la ganancia es limitada. Se parte de la hipótesis de que añadir como entradas al LSTM las salidas de modelos físicos junto con forzamientos meteorológicos (datos de temperaturas máxima y mínima, y precipitación) y atributos estáticos de la cuenca, debería mejorar la predicción de caudal al incorporar estados hidrológicos y restricciones dinámicas difíciles de inferir sólo con la meteorología. Resultados iniciales muestran una mejora frente a modelos físicos puros y un desempeño superior de LSTM regional frente a LSTM local. Además, la inclusión de parámetros estáticos aumenta la calidad del ajuste. No obstante, el valor añadido de introducir modelos físicos en el enfoque híbrido es pequeño en promedio, y se aprecia principalmente para caudales bajos, lo que motiva estudiar esta aparente “falta de mejora”. Para explicar este comportamiento se realizan dos pruebas: 1) una prueba de redundancia condicionada, donde se elimina la influencia meteorológica ajustando un modelo lineal (Ridge), tanto para el caudal observado como para la salida del modelo físico, y se evalúa la dependencia entre residuos; 2) una prueba de destilación de conocimiento, entrenando un LSTM para reproducir la salida de modelos físicos usando sólo forzamientos meteorológicos. Los resultados indican que el LSTM aprende a emular el modelo físico casi perfectamente, lo que sugiere que gran parte de la información aportada por el simulador no es independiente para la red, sino derivable a partir de los forzamientos meteorológicos. También se analiza la dependencia regional usando una clasificación de cuencas en cinco clústeres según sus características climáticas y físico-estáticas. En clústeres húmedos/templados se observa alta emulación y ganancia híbrida nula, en clústeres secos/transicionales el simulador puede introducir sesgo/ruido y en un clúster transicional con almacenamiento se identifican cuencas donde la modelización híbrida mejora de forma consistente. En general, los resultados evidencian una firma física predecible que permite anticipar en qué cuencas el enfoque híbrido aporta mejoras consistentes.

Agradecimientos: Esta investigación ha sido llevada a cabo en el marco del proyecto PID2021-126401OB-I00, financiado por MICIU/AEI/10.13039/501100011033 y por FEDER, UE.