

Simulación del caudal en España utilizando redes Long Short-Term Memory

Jesús Casado Rodríguez ¹

¹ Universidad de Alcalá de Henares; jesus.casador@edu.uah.es

* Correspondencia: jesus.casador@uah.es

Se ha creado el conjunto de datos CAMELS-ES dentro de la iniciativa CARAVAN para estudios hidrológicos a gran escala. Sobre estos datos se han entrenado dos redes neuronales recurrentes, una que predice el caudal observado, y otra que emula el funcionamiento del modelo hidrológico LISFLOOD-OS.

Resumen: Dentro de la iniciativa CARAVAN, que pretende crear un conjunto de datos hidrológicos globales, se ha creado el subconjunto de datos correspondiente a España. El punto de partida son las series temporales de caudal diario en ríos disponibles en el Anuario de Aforos de España. Tras el tratamiento y la selección de puntos de estudio, se ha generado un conjunto de datos con 269 cuencas. Los datos incluyen las series diarias de caudal, series meteorológicas de reanálisis extraídas de ERA5, y atributos estáticos (sin variación temporal) de las cuencas extraídos de HydroATLAS. Sobre estos datos se ha entrenado una red neuronal recurrente (Long-short Term Memory, LSTM) capaz de reproducir el caudal diario en el punto de salida de una cuenca a partir de las variables meteorológicas y los atributos de dicha cuenca. Además, CAMELS-ES cuenta con información adicional extraída de EFAS (European Flood Awareness System), que abarca series meteorológicas observadas extraídas de EMO1 (European Meteorological Observations), series de caudal simulado con el modelo hidrológico LISFLOOD-OS, un conjunto de atributos generados a partir de los mapas estáticos que utiliza LISFLOOD-OS, y los valores calibrados de los parámetros del modelo hidrológico. Con esta información extra se ha entrenado una segunda red LSTM que emula el funcionamiento del modelo hidrológico LISFLOOD-OS. El objetivo de esta segunda red es ser capaces de implementar en un futuro el paradigma de aprendizaje de parámetros en la calibración de LISFLOOD-OS.

Palabras clave: CAMELS-ES; CARAVAN; EFAS; LISFLOOD; LSTM

1. Introducción

Los modelos hidrológicos son algoritmos capaces de reproducir el ciclo del agua. Tradicionalmente, estos modelos están basados en ecuaciones que reproducen de forma aproximada los procesos físicos que tienen lugar en el ciclo hidrológico: precipitación, infiltración, escorrentía, flujo subterráneo del agua, traslación del caudal en los cauces... Los modelos se alimentan de datos meteorológicos (precipitación, temperatura, etc.) y fisiográficos (altitud, pendiente, características del suelo, de la vegetación, etc.) y simulan diversas variables de estado y flujos del ciclo hidrológico, siendo la más habitual el caudal, pero pudiendo abarcar otras variables como la cobertura nival, humedad del suelo, nivel freático, etc. Este tipo de modelos siguen siendo los más extendidos en el campo de la hidrología, por ejemplo, para la predicción de avenidas en tiempo real.

A este tipo de modelos pertenece LISFLOOD Open Source (Joint Research Centre - European Commission, 2023c; van der Knijff et al., 2010). LISFLOOD-OS se utiliza, entre otros, en las simulaciones de EFAS (European Flood Awareness System) (Joint Research



Centre - European Commission, 2023a) y de GloFAS (Global Flood Awareness System) (Joint Research Centre - European Commission, 2023b). Para su implementación en EFAS, el modelo LISFLOOD-OS fue calibrado en cerca de 2000 estaciones de aforo de toda Europa, partiendo de 76 mapas con las características físicas de las cuencas (Salamon et al., 2023) y las series temporales de caudal, precipitación y temperatura observadas (Thiemig et al., 2022). La calibración de este tipo de modelos se hace mediante algoritmos genéticos que requieren cientos o miles de simulaciones hasta optimizar una métrica de rendimiento que compara el caudal observado y el simulado (Grimaldi et al., 2023; Hirpa et al., 2018). La calibración es un proceso muy dilatado debido al coste computacional de simular un modelo físico, al alto número de iteraciones necesarias, y a que todo el proceso se ha de repetir en cada subcuenca hidrográfica independientemente. Este último punto es una de las principales carencias de este tipo de calibración. Por un lado, si no se hace una validación tanto espacial como temporalmente, el modelo podría adolecer de sobreajuste. Por otro lado, puesto que el algoritmo de calibración de una cuenca no conoce nada sobre la calibración de otras cuencas, no hay una coherencia espacial en los parámetros calibrados.

Durante los últimos años, con la irrupción de los modelos de aprendizaje automático y aprendizaje profundo, ha aparecido un nuevo tipo de modelos hidrológicos, los puramente basados en datos (Kratzert, Klotz, Brenner, et al., 2018; Nevo et al., 2022). Estos modelos utilizan la capacidad de las redes neuronales de extraer patrones de cantidades masivas de datos para simular aspectos del ciclo hidrológico, como puede ser el caudal de los ríos. En los últimos años se ha demostrado la potencialidad de las redes neuronales recurrentes, en concreto las LSTM (Long Short-Term Memory), para la simulación de series temporales de caudal (Feng et al., 2020; Gauch et al., 2021; Koch & Schneider, 2022; Kratzert et al., 2019; Kratzert, Klotz, Brenner, et al., 2018; Kratzert, Klotz, Herrnegger, et al., 2018; Lees et al., 2022; Mao et al., 2021; Yokoo et al., 2022). Existen estudios comparativos de estas LSTM y los modelos tradicionales donde se muestra que el uso de redes neuronales supera en rendimiento a los modelos tradicionales. Otras ventajas que se atribuyen al uso de LSTM son la rapidez tanto en la calibración del modelo (puesto que se calibra un único modelo para todas las cuencas de estudio, en lugar de una calibración para cada cuenca), y de ejecución (puesto que los modelos tradicionales o físicamente basados deben de resolver ecuaciones en ocasiones complicadas en mallas de cada vez mayor resolución). Por su parte, las desventajas que se atribuyen a este tipo de modelos son principalmente dos: el hecho de que son cajas negras donde no se sabe qué pasa dentro del modelo, y que su aplicación es específica a una o varias variables concretas (por ejemplo, el caudal), es decir, no se simulan el resto de los procesos del ciclo hidrológico.

Con la idea de extraer el potencial de ambos tipos de modelos surgen los modelos híbridos (Slater et al., 2023). En ellos se aplican redes neuronales dentro del proceso de modelado, ya sea en la calibración del modelo físico, en el pre-procesado, dentro del propio modelo, o en el post-procesado de los resultados. En el segundo de estos grupos, el que utiliza redes neuronales en la calibración, aparece el concepto de aprendizaje de parámetros (Tsai et al., 2021). La idea principal es que una red neuronal aprenda a estimar los parámetros del modelo físico en función de otras variables y así poder generar los parámetros adecuados para cualquier ubicación. Un requisito indispensable para poder aplicar aprendizaje de parámetros sobre un modelo físico es que sea diferenciable, de modo que pueda aplicarse el algoritmo de retropropagación en su entrenamiento. Un método alternativo es crear una red neuronal que emule el funcionamiento del modelo físico, lo que se conoce como un modelo subrogado. La técnica del aprendizaje de parámetros tiene dos ventajas principales. Por un lado, se calibra un único modelo que incluye todas las cuencas de interés. Por otro lado, como resultado se obtiene la relación entre las variables fisiográficas y los parámetros del modelo físico, lo que permite generar los parámetros del modelo físico en cuencas sin datos de caudal (lo que en hidrología se conoce como regionalización de parámetros).

A medio plazo, se quiere explorar la aplicabilidad del aprendizaje de parámetros en la calibración de LISFLOOD-OS de cara a su implementación en EFAS y GloFAS. Puesto

que LISFLOOD-OS no está escrito en un lenguaje diferenciable, es necesario, como paso previo, crear un modelo subrogado de LISFLOOD-OS.

El objetivo de este trabajo es triple. En primer lugar, crear un conjunto de datos hidrológicos para España que incluya los datos necesarios (atributos, series meteorológicas y series de caudal) de un conjunto amplio de cuencas que represente en lo posible los distintos regímenes hidrológicos en España. Este conjunto será publicado dentro de la comunidad hidrológica CARAVAN (Kratzert et al., 2023) bajo el nombre CAMELS-ES (Catchment Attributes and Meteorology for Large-sample Studies – España). En segundo lugar, desarrollar una red neuronal capaz de simular las series históricas de caudal en las cuencas de estudio. En último lugar, desarrollar una red neuronal que emule el funcionamiento del modelo hidrológico físicamente basado LISFLOOD-OS.

2. Materiales y métodos

2.1 Cuencas de estudio

El punto de partida para generar el conjunto de datos CAMELS-ES son las series diarias de caudal publicadas en el Anuario de Aforos (CEDEX, 2021). Esta base de datos contiene, entre otros, las series diarias de caudal de las estaciones de aforo dependientes del Ministerio de Transición Ecológica. Es decir, incluye datos de toda la España peninsular a excepción de las cuencas internas catalanas y andaluzas. En total, se incluyen 1074 estaciones con una representación geográfica diversa (ver Figura 2).

Sobre este conjunto inicial de datos es necesario filtrar aquellas estaciones de interés para este estudio. El filtrado de estaciones se hace en función de una serie de condiciones. Se seleccionan sólo estaciones con cuencas hidrográficas de al menos 100 km², criterio que se toma de CARAVAN. El periodo de interés abarca desde el 1 de octubre de 1991 (fecha de inicio de las simulaciones de EFAS5) hasta el 30 de septiembre de 2020 (fin de los datos en el Anuario de Aforos). Dentro de este periodo de estudio, se seleccionan aquellas estaciones con una cobertura de datos de al menos el 90% durante al menos 8 años consecutivos (criterio tomado de la calibración de EFAS5). Sobre las series así seleccionadas (606), se hace un filtrado manual para seleccionar aquellas estaciones (337) cuya serie de caudal no está claramente alterada por la presencia de embalses, vertidos u otras alteraciones humanas de complicada representación en un modelo hidrológico. Por último, para evitar la correlación espacial, se eliminan estaciones muy próximas en el mismo cauce.

2.2 CAMELS-ES

La iniciativa CARAVAN (Kratzert et al., 2023) incluye un repositorio GitHub con las herramientas necesarias para contribuir con un nuevo conjunto de datos. Los datos de partida son las series de caudal observado en estaciones de aforo (obtenidas en el punto anterior) y una capa GIS de polígonos con las cuencas hidrográficas de dichas estaciones. Para el trazado de las cuencas hidrográficas de las estaciones seleccionadas se utilizaron herramientas GIS (QGIS y ArcHydrotools de ArcMap), partiendo del modelo digital del terreno MERIT (Multi-Error-Removed Improved-Terrain DEM) (Yamazaki et al., 2017), de aproximadamente 90 m de resolución. Se seleccionó este modelo del terreno por estar ya corregido hidrológicamente de manera que el flujo de los ríos no se vea interrumpido por puentes, embalses u otros obstáculos.

Los códigos del repositorio CARAVAN utilizan la capa de polígonos para extraer tanto las series temporales de variables hidrometeorológicas, como los atributos de cada cuenca. Las series hidrometeorológicas se obtienen de las simulaciones del modelo meteorológico ERA5 (Hersbach et al., 2020), el modelo de reanálisis del ECMWF (European Centre for Medium-Range Forecast). Los atributos de las cuencas, a su vez, se extraen de la base de datos HydroATLAS (Lehner et al., 2022; Linke et al., 2019).

Junto con las series meteorológicas y los atributos de las cuencas, el conjunto de datos CAMELS-ES incluye las series de caudal observado seleccionadas a partir del Anuario de Aforos. Por coherencia de unidades con las series de precipitación y evapotranspiración

(en mm) las series de caudal en CAMELS-ES son en realidad de caudal específico (mm/d), para lo que se divide la serie original por el área de la cuenca vertiente.

2.3 Ampliación de CAMELS-ES: EMO1 y EFAS5

El conjunto de datos CAMELS-ES básico, el que se genera a partir de las herramientas del repositorio de CARAVAN, ha sido ampliado con los datos provenientes del sistema EFAS en su reciente versión 5.

EFAS es un sistema de pronóstico hidrológico basado en la simulación del modelo hidrológico LISFLOOD-OS (Joint Research Centre - European Commission, 2023c). LISFLOOD-OS simula el ciclo hidrológico a partir de las series meteorológicas de entrada (precipitación, temperatura y evapotranspiración) del conjunto de datos EMO1 (Thiemig et al., 2022), una serie de 76 mapas estáticos con características de las cuencas (geomorfología, usos del suelo, vegetación, características del suelo, demanda...) (Salamon et al., 2023) y los 14 parámetros del modelo calibrados para cada una de las cuencas.

Toda la información del sistema EFAS se agregó espacialmente y se añadió al conjunto de datos CAMELS-ES. A partir de los mapas de EMO1, se calcularon las series temporales areales de precipitación, temperatura y evapotranspiración de cada cuenca. A partir de los mapas estáticos se calcularon diversos estadísticos (media, suma, desviación típica, mínimo, máximo) de las características de las cuencas y se agregaron a los atributos de CAMELS-ES. De forma similar, se calculó la media de los parámetros del modelo en cada una de las subcuencas y se agregaron como atributos en CAMELS-ES. Por último, se extrajo la serie de caudal simulado (reanálisis) de EFAS5 en los puntos de salida de cada una de las cuencas, es decir, en las celdas del modelo que representan la estación del Anuario de Aforos.

El objetivo de esta ampliación es doble. Por un lado, dotar a futuros usuarios del conjunto de datos con más información sobre las características de las cuencas y un hito de comparación del rendimiento de su modelo. Por otro, disponer en el formato adecuado la información de EFAS para poder entrenar una red neuronal que emule el funcionamiento del modelo hidrológico LISFLOOD-OS.

2.4 Simulación del caudal con LSTM

Como se describió en la introducción, hay numerosos trabajos en los últimos años que utilizan redes neuronales recurrentes en la simulación del caudal en ríos, en su mayoría redes de tipo LSTM. La ventaja de las redes LSTM en hidrología es que el estado de la celda, su memoria, permite emular el funcionamiento de una cuenca hidrológica, donde su respuesta en forma de caudal está condicionada por situaciones a distintas escalas del pasado que afectan a los diversos almacenamientos de agua en la cuenca (humedad del suelo, nivel inicial de los ríos, agua almacenada en forma de nieve...).

La librería Neural Hydrology (Kratzert et al., 2022) es una compilación de las múltiples herramientas utilizadas en diversos artículos de investigación sobre el potencial de las redes LSTM en hidrología. Está construida sobre PyTorch, una de las librerías de redes neuronales de más difusión en el lenguaje Python y desarrollada inicialmente por Facebook. Neural Hydrology contiene una gama de modelos LSTM de diversa índole y con diversos fines. En este trabajo se utiliza el modelo *cuda_lstm*, una de las implementaciones más sencillas que permite la simulación de un conjunto de series temporales a partir de datos dinámicos de entrada (series meteorológicas, por ejemplo) y/o datos estáticos (atributos de la cuenca). La red se alimenta de los datos dinámicos de una ventana de tiempo del pasado y es capaz de predecir el valor de la variable objetivo (en este caso caudal) en el paso temporal posterior. En caso de incluirse datos estáticos, se concatenan a cada uno de los pasos temporales de esa ventana de tiempo pasado. Repitiendo este proceso de simulación en una ventana móvil se simula la serie temporal objetivo.

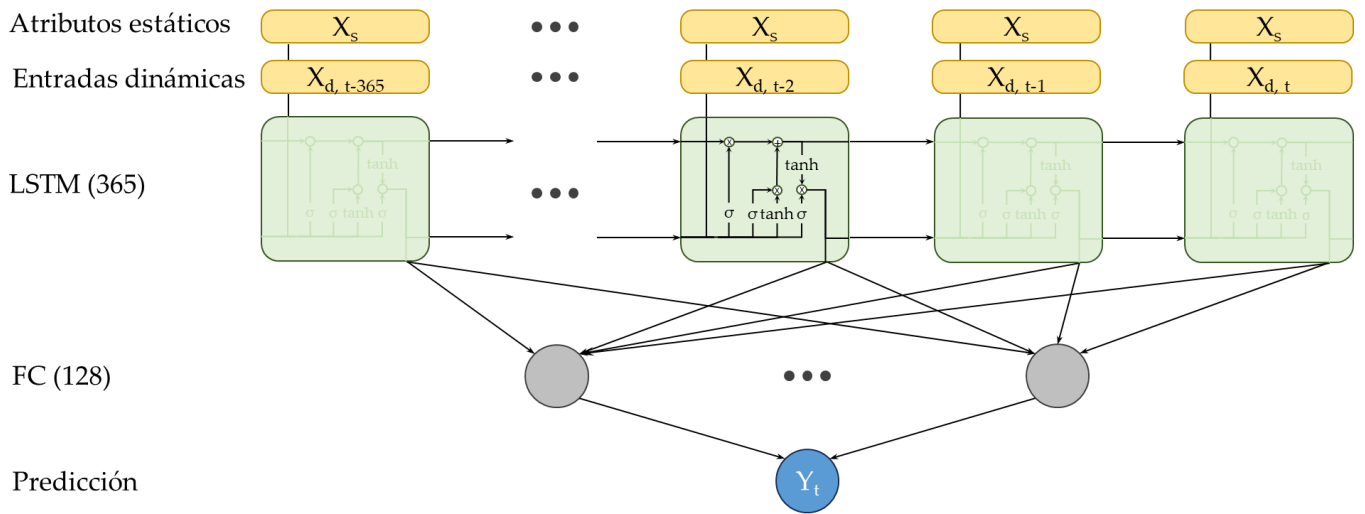


Figura 1. Esquema de la red neuronal utilizada para predecir un dato diario de caudal (Y_t) a partir de unos atributos estáticos de las cuencas (X_s) y las series temporales de los últimos 365 días (X_d). La red se compone de una red neuronal recurrente de tipo LSTM unida a una capa completamente conectada.

La Figura 1 muestra el esquema general de la red neuronal utilizada para simular caudal. La red se compone de dos capas únicamente. Una primera capa de LSTM de tantos nodos como el tamaño de la ventana de tiempo pasado utilizada en la predicción. La salida de la red recurrente pasa por una capa totalmente conectada con activación lineal, cuyo número de nodos fue uno de los parámetros a ajustar.

Siguiendo las recomendaciones la bibliografía, se entrenan redes LSTM de 365 nodos que utilizan las series de entrada de los últimos 365 días para simular el caudal del día siguiente. Además, se hicieron múltiples simulaciones para obtener los valores adecuados del número de nodos de la red conectada (128), el dropout (0.4), el batch size (256) y el límite de los gradientes normalizados (1). Como algoritmo de optimización se utiliza Adam (Adaptive Moment Estimation), siendo la función de coste el coeficiente de eficiencia de Nash-Sutcliffe (NSE) (Nash & Sutcliffe, 1970), una métrica basada en errores cuadráticos ampliamente usada en hidrología (ec. 1).

$$NSE = \frac{\sum_{t=1}^T (Q_{obs} - Q_{sim})^2}{\sum_{t=1}^T (Q_{obs} - \bar{Q}_{obs})^2} \quad (1)$$

Donde Q_{obs} y Q_{sim} son, respectivamente, el caudal observado y simulado en un paso de tiempo t , y \bar{Q}_{obs} el caudal medio observado. El NSE toma valores desde menos infinito a 1, donde 1 es el valor óptimo y valores negativos representan modelos peores que un modelo cuya predicción fuera simplemente el caudal medio. Para la validación del modelo y su comparación con el rendimiento de EFAS5 se utiliza, sin embargo, el coeficiente de eficiencia de Kling-Gupta (Gupta et al., 2009), otra métrica de rendimiento ampliamente utilizada en hidrología que está sustituyendo al NSE. La métrica está basada en la distancia euclídea de tres componentes (coeficiente de correlación, el cociente de las desviaciones típicas, y el cociente de las medias) con respecto a 1, su valor óptimo. Al igual que el NSE, el valor óptimo del KGE es 1, pero en este caso un modelo basado en el caudal medio tiene un KGE de -0.41, no de 0.

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \quad (2)$$

Se han tenido que utilizar distintas métricas en el entrenamiento y la validación porque la librería Neural Hydrology no tiene implementado KGE como función de coste para el entrenamiento.

2.4.1 Simulación del caudal observado con LSTM

Esta red corresponde al segundo de los objetivos de este estudio: crear un modelo LSTM capaz de simular el caudal diario en cualquier cuenca de España. La variable objetivo son las series de caudal específico diario observadas en las estaciones del Anuario de Aforos, y las variables predictoras un subconjunto de variables estáticas (Tabla 1) y dinámicas (Tabla 2) de CAMELS-ES.

En hidrología, donde los datos tienen una componente tanto espacial como temporal, los datos de partida se han dividido en estas dos dimensiones con el objetivo de asegurar que el modelo funciona adecuadamente en otras cuencas (validación espacial) y en periodos distintos (validación temporal) al de entrenamiento. Por un lado, las cuencas de estudio se dividen en tres subconjuntos de entrenamiento (60%), validación (20%) y evaluación (20%). Este muestreo de estaciones se hace por demarcaciones hidrográficas para evitar que las cuencas del Sur y Este, con menos estaciones, no estén representadas en los datos de entrenamiento. Por otro lado, los periodos de estudio se han definido específicamente para cada estación, evitando en lo posible que el periodo de entrenamiento coincida con el periodo de validación. Para cada estación, se parte de la serie de años consecutivos más larga de su historia; en las estaciones de entrenamiento se utiliza únicamente el 60% de datos más recientes (bajo la suposición de que los datos más recientes son más fiables), mientras que en las estaciones de validación se utiliza el 60% de datos más antiguos. En las estaciones de evaluación se utiliza toda la serie temporal disponible.

Tabla 1. Atributos estáticos de CAMELS-ES utilizados en el modelo LSTM de caudal observado.

Categoría	Atributo	Descripción
Topografía	ele_mt_sav	Altitud media (m)
	ele_mt_smn	Altitud mínima (m)
	ele_mt_smx	Altitud máxima (m)
	slp_dg_sav	Pendiente media (°)
Usos del suelo	for_pc_sse	Fracción de bosque
	crp_pc_sse	Fracción de cultivos
	ire_pc_sse	Fracción de regadío
	lka_pc_sse	Fracción de lagos
	pst_pc_sse	Fracción pasto
Tipos de suelo	snd_pc_sav	Fracción de arena
	cly_pc_sav	Fracción de arcilla
	kar_pc_sse	Fracción de roca carbonatada
	swc_pc_syr	Máximo contenido de humedad del suelo
	snow_pc_syr	Fracción de nieve
Hidrología	p_mean	Precipitación media (mm)
	pet_mean	Evapotranspiración potencial media (mm)
	aridity	Aridez
	moisture_index	Índice de humedad
	seasonality	Estacionalidad
	high_prec_freq	Frecuencia de precipitación intensa
	high_prec_dur	Duración de precipitación intensa
	low_prec_freq	Frecuencia de precipitación débil
	low_prec_dur	Duración de precipitación débil

Tabla 2. Atributos dinámicos de CAMELS-ES utilizados en el modelo LSTM de caudal observado.

Categoría	Atributo	Descripción
Meteorología	total_precipitation_sum	Precipitación total diaria (mm)
	temperature_2m_min	Temperatura mínima diaria (°C)
	temperature_2m_max	Temperatura máxima diaria (°C)
	potential_evaporation_sum	Evaporación potencial diaria (mm)
	surface_net_solar_radiation_mean	Radiación solar neta media (W/m ²)
	surface_pressure_mean	Presión atmosférica media (kPa)
Variable objetivo streamflow		Caudal específico (mm/d)

2.4.2 Emulación del modelo LISFLOOD-OS con LSTM

Este apartado corresponde al objetivo 3 de este estudio: crear un modelo LSTM capaz de emular al modelo hidrológico LISFLOOD-OS. Para ello se utilizan como datos de entrada los utilizados en el sistema EFAS que fueron añadidos como extensión en CAMELS-ES. Es decir, como variables dinámicas se utilizan las series de precipitación, temperatura y evapotranspiración de EMO1 (Tabla 3), el conjunto de datos climáticos utilizados en la calibración de EFAS. Como variables estáticas se utilizan los atributos de las cuencas generados a partir de los mapas estáticos y los parámetros calibrados de LISFLOOD (Tabla 4). La variable objetivo siguen siendo las series de caudal específico en las estaciones, pero en este caso se utilizan las series simuladas en EFAS, no las series observadas.

La selección de estaciones de entrenamiento, validación y evaluación es idéntica a la del primer modelo LSTM, no así los periodos de estudio. Como las series de EFAS abarcan todo el periodo desde 1991 hasta 2020, se puede definir un mismo periodo de estudio para todas las estaciones siguiendo la misma lógica del primer modelo. Es decir, en el entrenamiento se utiliza el 60% más reciente de la serie, en la validación el 60% más antiguo y en la evaluación la serie completa.

3. Resultados

3.1 CAMELS-ES

La Figura 2 muestra la distribución geográfica de las estaciones del Anuario de Aforos seleccionadas para formar parte de CAMELS-ES, así como las descartadas. El sombreado oscuro representa el área que cubren las cuencas hidrográficas de CAMELS-ES. El histograma de la izquierda muestra el número de estaciones que originalmente contiene el Anuario de Aforos en cada demarcación hidrográfica y las seleccionadas.

La muestra final incluye 269 cuencas que se concentran principalmente en el Centro y Norte peninsular. En las cuencas del Sur (Guadiana y Guadalquivir) y Sureste (Júcar y Segura) el grado de alteración de las cuencas es muy alto, por lo que la densidad de puntos seleccionados es muy baja. Es de notar que sólo una estación del Guadiana ha sido seleccionada a pesar de que esta demarcación es la cuarta por número de estaciones. En el Segura, por su parte, no se ha seleccionado ninguna de las más de 50 estaciones de las que dispone.

Tabla 3. Atributos dinámicos de EFAS utilizados en el modelo LSTM que emula LISFLOOD-OS.

Categoría	Atributo	Descripción
Variables de entrada	pr_emo1	Precipitación total diaria (mm)
	ta_emo1	Temperatura media diaria (°C)
	e0_emo1	Evapotranspiración diaria de referencia (mm)
Variable objetivo	dis_efas5	Caudal específico (mm/d)

Tabla 4. Atributos estáticos de EFAS5 utilizados en el modelo LSTM que emula LISFLOOD-OS.

291

Categoría	Atributo	Descripción
Topografía	ele_mean	Altitud media (m)
	ele_min	Altitud mínima (m)
	ele_max	Altitud máxima (m)
	gradient_mean	Pendiente media (°)
	uparea_max	Área de la cuenca (m ²)
Usos del suelo	forest	Fracción de bosque
	irrigated	Fracción de regadío
	other	Fracción de otros usos
	rice	Fracción de arrozal
	sealed	Fracción de suelo urbano
	water	Fracción de cuerpos de agua
Vegetación	laiymax	Índice foliar máximo
	laiyrmmean	Índice foliar medio
	laiyrmin	Índice foliar mínimo
	cropcoef_mean	Coefficiente de cultivo medio
Características del suelo	ksat1	Conductividad hidráulica saturada capa 1 (cm/d)
	lambda1	Índice del tamaño de los poros
	genua1	Coefficiente de Van Genuchten
	soildepth1	Profundidad del suelo (mm)
	thetas1	Humedad del suelo saturada
	thetar1	Humedad del suelo residual
Cuerpos de agua	reserv_vol	Volumen total de embalses (m ³)
	reserv_count	Número de embalses
	lake_area	Área total de lagos (m ²)
	lake_count	Número de lagos
Demandas	dom_yr	Demanda doméstica anual (mm)
	ene_yr	Demanda energética anual (mm)
	ind_yr	Demanda industrial anual (mm)
	liv_yr	Demanda ganadera anual (mm)
Parámetros de LISFLOOD	adjust_normal_flood	Ajuste del nivel de avenida de un embalse
	b_xinjiang	Coef. de infiltración de Xinjiang
	calchanman1	Coef. de Manning del canal de aguas bajas
	canchanman2	Coef. de Manning del canal de aguas altas
	gwloss	Pérdidas del acuífero profundo (mm/d)
	gwpercvalue	Percolación máxima (mm/d)
	lakemultiplier	Ajuste del caudal de salida de los lagos
	lowerzonetimeconstant	Tiempo de residencia del acuífero inferior (d)
	lzthreshold	Umbral para el caudal base (mm)
	powerprefflow	Exponente del flujo preferencial
	qsplitmilt	Límite entre caudal de aguas bajas y altas
	reservoirnormqmult	Ajuste del caudal normal de salida del embalse
	snowmeltcoef	Coefficiente de derretimiento (mm·°C ⁻¹ ·d ⁻¹)
	upperzonetimeconstant	Tiempo de residencia del acuífero superior (d)

292

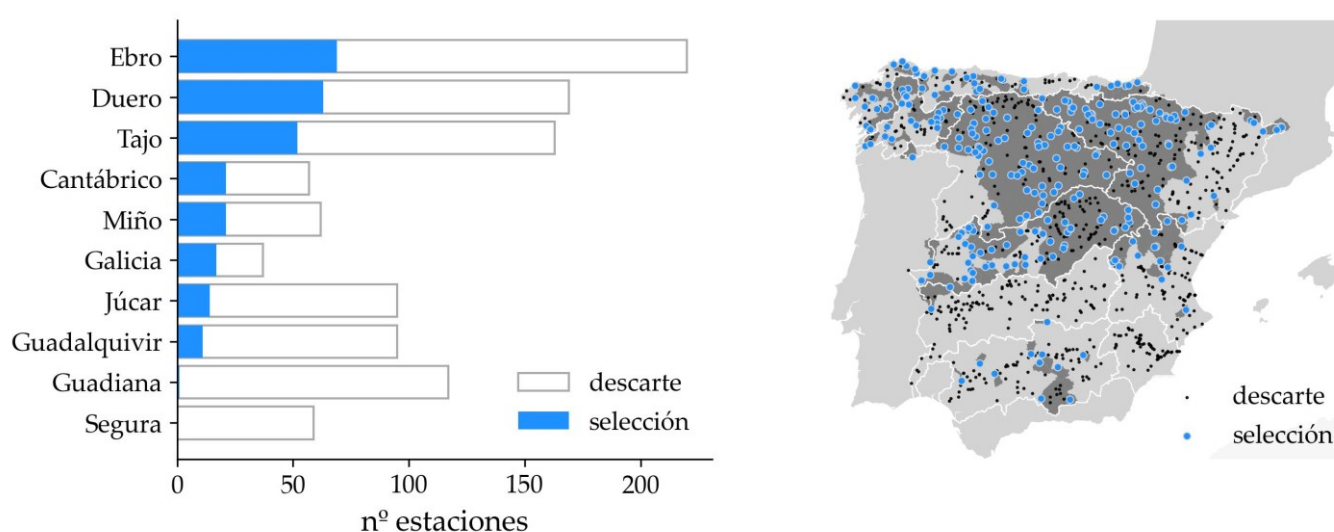


Figura 2. Estaciones del Anuario de Aforos y subconjunto seleccionado para su inclusión en CAMELS-ES. El área sombreada representa las cuencas hidrográficas de dichas estaciones seleccionadas. Las líneas blancas representan la divisoria de las demarcaciones hidrográficas.

3.2 Simulación del caudal

3.2.1 Rendimiento de EFAS

Como hito de referencia, la Figura 3 muestra el rendimiento en las cuencas de CAMELS-ES en el modelo hidrológico LISFLOOD-OS según su implementación en EFAS5. Se muestra tanto la función de distribución acumulada del KGE de las 269 cuencas, como su distribución geográfica.

Se compararon las cuencas calibradas en EFAS5 con las seleccionadas para CAMELS-ES con la idea de que ambos conjuntos fueran coherentes. Sin embargo, no se pudo conseguir puesto que buena parte de las estaciones calibradas en EFAS5 se descartaron de CAMELS-ES por tener series de poca calidad, mientras que cuencas seleccionadas en CAMELS-ES no fueron calibradas en EFAS5. Esto se resume en que el rendimiento de EFAS5 que muestra la Figura 3 incluye tanto cuencas calibradas como no.

El rendimiento de EFAS en España es relativamente pobre (KGE mediano de 0.383). Es un problema conocido y que no es representativo del rendimiento de EFAS en todo el dominio europeo. Este pobre rendimiento está causado por dos motivos principalmente. Los modelos hidrológicos rinden peor en climas áridos como el de buena parte de España. Además, la alta densidad de embalses en España y su representación simplista en LISFLOOD (como en el resto de los modelos hidrológicos) impide reproducir adecuadamente regímenes de caudal fuertemente alterados. Ambos motivos quedan reflejados en el mapa de la Figura 3, donde las cuencas con peor rendimiento son de la mitad sur peninsular, de clima árido, o son cuencas de cabecera del Duero y Ebro, habitualmente reguladas por embalses.

Se ha de remarcar en este punto que el segundo modelo LSTM va a intentar reproducir estas series de caudal, no las observadas, con lo que aprenderá a simular series de por sí erróneas en muchos casos. Esto puede carecer de sentido a primera vista, pero el objetivo de ese segundo modelo LSTM no es simular adecuadamente el ciclo hidrológico, sino explorar la posibilidad de crear una red neuronal capaz de emular LISFLOOD de cara a poder cambiar en el futuro el método utilizado en su calibración.

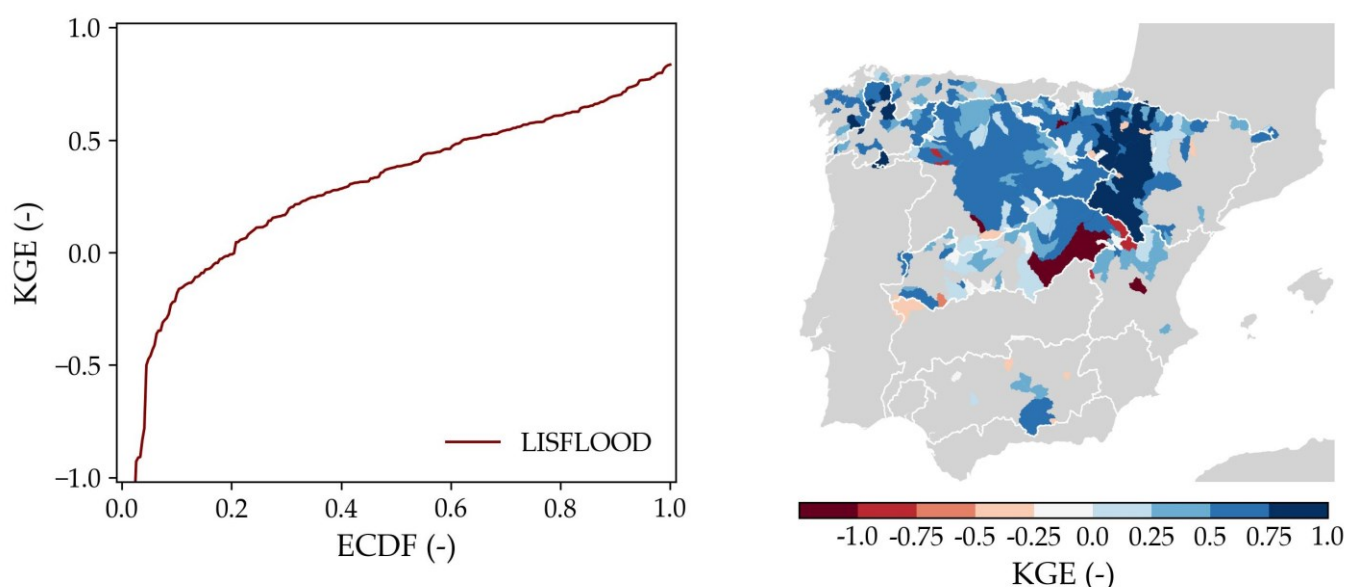


Figura 3. Rendimiento (medido en KGE) del modelo hidrológico LISFLOOD según su implementación en EFASv5. En el panel izquierdo, la función de densidad empírica del KGE. En el panel derecho, la distribución geográfica del KGE.

3.1 Simulación del caudal observado con LSTM

El rendimiento del modelo LSTM que simula las series de caudal observado del Anuario de Aforos se resume en la Figura 4. En este caso se muestran por separado las funciones de distribución de las muestras de entrenamiento, validación y evaluación, y como referencia el rendimiento de EFAS5 ya mostrado en el apartado anterior.

El rendimiento de la red neuronal es netamente superior al de EFAS en las cuencas de CAMELS-ES (KGE mediano de 0.693). Esta comparativa es aún más favorable hacia el LSTM cuando se tiene en cuenta que el rendimiento de EFAS aquí mostrado incluye tanto cuencas calibradas como no, y aun así es superado por las muestras de validación y test del LSTM. El modelo tiende al sobreajuste, con un KGE mediano de 0.730, 0.593, 0.570 en las muestras de entrenamiento, validación y evaluación, respectivamente. Se entrenaron modelos con diversos hiperparámetros para limitar dicho sobreajuste, optando finalmente por un *dropout* de 0.4 y un valor del límite de los gradientes normalizados de 1.

A pesar de que el rendimiento de la mayoría de las cuencas es bueno (el 77% tiene un KGE superior a 0.5), existen aún una serie de cuencas con un rendimiento muy pobre. El mapa muestra un claro gradiente de rendimiento desde el Sureste hacia el Noroeste peninsular. Destaca el alto rendimiento de las cuencas de Galicia Costa y Miño-Sil, las más húmedas de la Península, y la notable mejoría del rendimiento con respecto a EFAS en el Duero y Ebro.

Buena parte de las cuencas con un rendimiento pobre se encuentran en el Júcar. El hecho de que LISFLOOD tuviera un mejor rendimiento en esta demarcación denota que puede haber algún proceso hidrológico de especial importancia en esta demarcación que la red neuronal no ha podido captar. Siendo la única demarcación de clima estrictamente mediterráneo y teniendo una representatividad mínima dentro del global del conjunto de datos (Figura 2), podría ser que el modelo no haya podido aprender a reproducir procesos en cuencas mediterráneas, con regímenes hidrológicos intermitentes y torrenciales. Otro posible motivo para este pobre rendimiento del Júcar sería la fuerte regulación causada por los embalses, algo que también existe en el Tago, Duero y Ebro, y que podría explicar el pobre rendimiento de ciertas cuencas de cabecera en el Ebro y Tago.

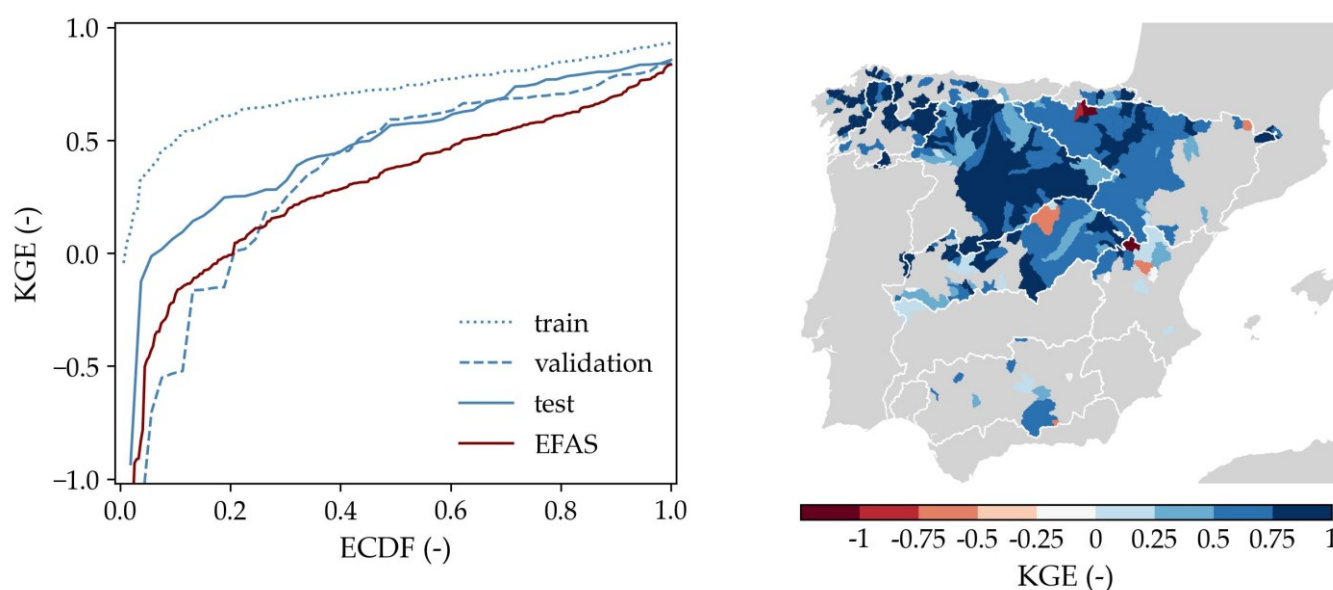


Figura 4. Rendimiento (medido en KGE) del modelo LSTM basado en el conjunto de datos CAMELS-ES. En el panel izquierdo, la función de densidad empírica del KGE en las muestras de entrenamiento, validación y evaluación, y la de EFAS5 como referencia. En el panel derecho, la distribución geográfica del KGE de todas las cuencas simuladas.

3.2 Emulador del modelo LISFLOOD-OS

La Figura 5 muestra el rendimiento del modelo LSTM que emula el caudal simulado en EFAS5. De nuevo la función de distribución muestra por separado el rendimiento de las muestras de entrenamiento, validación y evaluación, mientras que el mapa muestra todas las cuencas de forma conjunto.

El rendimiento es en general alto: KGE mediano de 0.732 en la muestra de evaluación. Este rendimiento era esperable, puesto que la tarea de reproducir el caudal simulado por un modelo hidrológico es más sencilla que la de reproducir el caudal real. En este ejercicio se han eliminado tanto las incertidumbres asociadas a los datos de entrada (errores en las mediciones meteorológicas y de caudal) como la incertidumbre inherente al modelo hidrológico (que puede no representar todos los procesos que acontecen en una cuenca). En este caso se sabe que con los datos meteorológicos de EMO1, una serie de ecuaciones fijas para todas las cuencas, y unos parámetros del modelo específicos de cada cuenca se consigue el caudal objetivo. La idea es que una red neuronal sea capaz de aprender la relación entre esas tres componentes (meteorología, ecuaciones y parámetros) de forma general para todas las cuencas de CAMELS-ES. Por ello, se podría decir que el rendimiento no es el esperado.

Al igual que en apartados anteriores, se aprecia un gradiente en el rendimiento de Sureste a Noroeste. Si bien en este caso las cuencas del Sureste tienen KGE por encima de 0, destacan claramente por su alto rendimiento las cuencas del Miño-Sil, Galicia Costa y Cantábrico. Estas cuencas se caracterizan por ser más pequeñas, más húmedas y menos reguladas. El área de la cuenca puede ser un factor influyente en el rendimiento del modelo; las cuencas grandes (en Duero, Ebro y Tajo) obtienen rendimientos medios, mientras que en el modelo de la sección anterior tenían rendimientos altos. Sin embargo, no hay una correlación clara entre área de cuenca y rendimiento porque hay mucha dispersión en el rendimiento de las cuencas pequeñas; tanto los mejores como los peores rendimientos se obtienen en cuencas de pequeño tamaño.

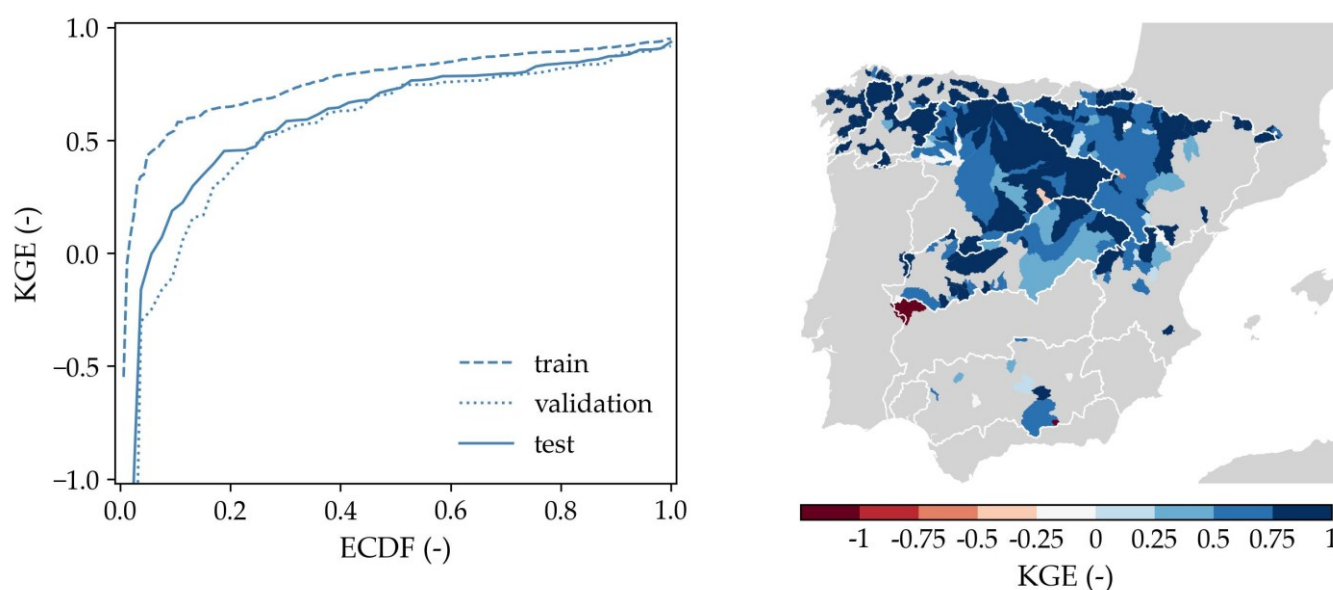


Figura 5. Rendimiento (medido en KGE) del modelo LSTM que emula la implementación del modelo hidrológico LISFLOOD en el sistema EFASv5. En el panel izquierdo, la función de densidad empírica del KGE en las muestras de entrenamiento, validación y evaluación. En el panel derecho, la distribución geográfica del KGE de todas las cuencas simuladas.

Los problemas de rendimiento de ciertas cuencas en el LSTM se pueden deber a dos factores. Por un lado, a que su parametrización en EFAS no es coherente con el global de las cuencas. Como se explicó en la introducción, ésta es una de las limitaciones de las calibraciones habituales en hidrología, donde no hay mecanismos que transfieran información sobre la parametrización entre las cuencas. Una segunda causa podría ser que el modelo LSTM no ingiere tanta información como LISFLOOD. LISFLOOD es un modelo distribuido que recibe mallas de relativamente alta resolución (aprox. 1.4 km) tanto de los atributos como de los forzamientos meteorológicos y resuelve las ecuaciones en cada celda de la malla. El LSTM, en cambio, recibe atributos y datos meteorológicos agregados a nivel de cuenca, con lo que la cantidad de información es mucho menor. Es decir, el modelo LSTM pierde la variabilidad espacial propia de un modelo de la Tierra. Ahondando más en este punto, LISFLOOD conoce la ubicación y tamaño de cada lago y embalse, mientras que el LSTM únicamente conoce el número y tamaño global de los lagos y embalses de cada cuenca. A riesgo de ser repetitivo, la presencia de embalses en las cuencas es un factor claramente influyente en su rendimiento, puesto que depende de un factor humano difícil de generalizar en un modelo.

4. Discusión

Este estudio presenta la primera versión del conjunto de datos CAMELS-ES. Este conjunto de datos se ha publicado en abierto en Zenodo y se incluirá dentro del conjunto de datos hidrológicos globales CARAVAN, lo que permitirá a científicos de todo el mundo disponer de datos de caudal en España que, a pesar de ser ya públicos, pueden ser desconocidos a nivel internacional.

Esta primera versión es susceptible de ser mejorada en el futuro. En primer lugar, la cobertura de las cuencas incluidas no es equitativa dentro del país. La mitad Sur de la Península está claramente infrarrepresentada debido a que las series de caudal del Anuario de Aforos para esta área eran de mala calidad. En un futuro se buscarán nuevas fuentes de datos como la red SAIH (Sistema Automático de Información Hidrológica) de cada demarcación o el sistema SIAR (Sistema de Información Agroclimática para el Regadío) para conseguir una mayor representación de las cuencas del Sur peninsular. Además, por no estar incluidas en el Anuario de Aforos, no hay ninguna estación en las cuencas

mediterráneas andaluzas y catalanas. En un futuro se buscará dicha información en la Junta de Andalucía y la Agencia Catalana del Agua. Por lo general, estas fuentes de datos alternativas contienen información a una resolución temporal mayor de la diaria (del orden de minutos), con lo que CAMELS-ES se podría aplicar a otro tipo de estudios hidrológicos como el pronóstico de avenidas. Por ejemplo, la comparación con EFAS sería más justa si los datos tuvieran resolución 6 h, el paso temporal de las simulaciones de EFAS.

Conocida la dificultad de simular regímenes hidrológicos fuertemente alterados por la presencia de infraestructuras hidráulicas, la primera versión de CAMELS-ES ha descartado tanto las series de caudal de dichas cuencas como las series propias de los embalses, que también incluye el Anuario de Aforos. Vista la importancia que está cobrando en el mundo de los modelos hidrológicos globales la mejora en la simulación de embalses (Hanazaki et al., 2022; Steyaert et al., 2022; Turner et al., 2021), sería de interés que futuras versiones de CAMELS-ES incluyeran series con las operaciones de los embalses.

Las dos redes LSTM entrenadas en este estudio son sólo una primera aproximación al uso de LSTM en la simulación hidrológica. La red LSTM entrenada con las series de caudal observado del Anuario de Aforos ha confirmado el potencial de este tipo de modelos, puesto que claramente supera en rendimiento al sistema EFAS. Con todo, se debería seguir trabajando sobre este modelo para mejorar su generalización en cuencas áridas o influenciadas por la presencia de embalses. El primero de estos problemas podría aliviarse con la inclusión de un mayor número de cuencas áridas en CAMELS-ES, con lo que la red tendría una mayor muestra de la que aprender los procesos hidrológicos en estos climas. Respecto a los embalses, existen en la literatura aplicaciones de aprendizaje profundo para aprender reglas generales de funcionamiento de los embalses (Liu et al., 2019), pero se requiere mayor investigación sobre cómo embeber la simulación de embalses dentro de un modelo más general. En este estudio se ha utilizado un conjunto de atributos estáticos genérico, pero sería de interés hacer un análisis de sensibilidad para limitar los datos de entrada a aquellos atributos que verdaderamente tienen influencia sobre la simulación.

La segunda red LSTM se ha entrenado como un modelo subrogado de LISFLOOD-OS. A pesar de que el rendimiento en la simulación de caudal es aceptablemente bueno, esta red requiere de mejoras para poder sustituir a LISFLOOD en una calibración que utilice el paradigma de aprendizaje de parámetros. Los resultados muestran que la capacidad de generalización del modelo no es buena, puesto que se observa una dependencia entre el rendimiento del modelo y la posición geográfica, clima y área de la cuenca. Además, un modelo subrogado de LISFLOOD debería poder replicar no sólo las series de caudal simulado, sino otros flujos y variables de estado del modelo como la evapotranspiración o la humedad del suelo, para lo que sería de especial interés entrenar la red LSTM frente a múltiples series temporales, no sólo el caudal.

El uso de las redes LSTM en la simulación hidrológica está cobrando mucha fuerza dado el rendimiento que han demostrado, pero todos los estudios hasta la fecha se centran en la simulación agregada. Mientras que la hidrología basada en procesos ha ido abandonando los modelos agregados (a escala cuenca) y adoptando modelos distribuidos (en mallas regulares) de cada vez mayor resolución capaces de reproducir la heterogeneidad propia de un sistema complejo como la Tierra, los modelos LSTM están aún anclados en el paradigma de simulación agregada. Parece intuitivo que el uso de redes convolucionales debería ser una vía para crear *embeddings* que contengan la información espacial original de los mapas de atributos y forzamientos meteorológicos (Kraft et al., 2022), pero su aplicación en hidrología se complica debido al diverso tamaño y geometría de las cuencas y a la necesidad de agregar el caudal según las direcciones de flujo de la red fluvial.

5. Conclusiones

En este estudio se ha generado un nuevo conjunto de datos hidrológicos de España disponible para su uso en estudios hidrológicos a gran escala. El conjunto, llamado CAMELS-ES, se enmarca dentro de la iniciativa CARAVAN para crear un conjunto global de datos hidrológicos que se utilice como referencia en el desarrollo de modelos a escala

global. La información que incluye CAMELS-ES no es más que una recopilación y agregación a escala de cuenca de datos públicos, tanto meteorológicos, atributos físicos como series de caudal observado. A los datos habituales en otros conjuntos de datos de CARAVAN, en CAMELS-ES se ha añadido además información proveniente del sistema EFAS relativa a meteorología observada, atributos físicos y caudal simulado. Se espera que la publicación de CAMELS-ES induzca a la aparición de España en estudios hidrológicos globales.

El conjunto de datos CAMELS-ES se ha utilizado para entrenar dos redes LSTM. Una primera red se ha entrenado para simular el caudal observado y comparar su rendimiento con el del sistema EFAS. Los resultados aquí mostrados confirman el potencial de las redes LSTM en la simulación hidrológica ya reportado en numerosa bibliografía científica. La segunda red se ha entrenado con la idea de crear un modelo subrogado de LISFLOOD-OS susceptible de ser utilizado en futuras calibraciones de EFAS. Los resultados muestran que este modelo subrogado requiere aún de más trabajo para asegurar un rendimiento similar en cuencas de diverso tipo e incluir un entrenamiento multivariable para que la red neuronal no sólo reproduzca el caudal simulado por LISFLOOD sino otras variables hidrológicas de importancia como la evapotranspiración, la humedad del suelo o la cobertura nival.

Disponibilidad de los datos: El conjunto de datos hidrológicos CAMELS-ES está disponible en Zenodo (<https://doi.org/10.5281/zenodo.8394799>) y será añadido tras su revisión a la iniciativa CARAVAN. Los códigos utilizados para el desarrollo de este trabajo son accesibles en el repositorio GitHub (https://github.com/casadoj/CAMELS-ES_LSTM).

Reconocimientos: Quisiera reconocer el trabajo del CEDEX y las Confederaciones Hidrográficas, que mantienen las redes de seguimiento hidrológicas y ponen a disposición del público los datos recopilados; espero que CAMELS-ES haga su trabajo más accesible a la comunidad científica. Agradecer a mis compañeros Peter Salamon y Stefania Grimaldi su apoyo en el diseño del estudio, y a mi compañero Gonçalo Gomes Ramos en la extracción de los datos brutos de EMO1 y EFAS5.

Referencias

- CEDEX. (2021). *Anuario de aforos 2018-2019*.
- Feng, D., Fang, K., & Shen, C. (2020). Enhancing Streamflow Forecast and Extracting Insights Using Long-Short Term Memory Networks With Data Integration at Continental Scales. *Water Resources Research*, 56(9). <https://doi.org/10.1029/2019WR026793>
- 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>
- Grimaldi, S., Salamon, P., Russo, C., Disperati, J., Zsoster, E., Carton de Wiart, C., Mazzetti, C., Choulga Margarita, Moschini, F., Harrigan, S., Gomes, G., Casado-Rodríguez, J., Ramos, A., Barnard, C., Hansford, E., & Prudhomme, C. (2023). GloFAS v4.0: towards hyper-resolution hydrological modelling at global scale. *European Geoscience Union General Assembly 2023*.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2), 80-91. <https://doi.org/10.1016/J.JHYDROL.2009.08.003>
- Hanazaki, R., Yamazaki, D., & Yoshimura, K. (2022). Development of a Reservoir Flood Control Scheme for Global Flood Models. *Journal of Advances in Modeling Earth Systems*, 14(3). <https://doi.org/10.1029/2021MS002944>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita,

- M., ... Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049. <https://doi.org/10.1002/qj.3803>
- Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., & Dadson, S. J. (2018). Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. *Journal of Hydrology*, 566, 595-606. <https://doi.org/10.1016/j.jhydrol.2018.09.052>
- Joint Research Centre - European Commission. (2023a). *European Flood Awareness System*. <https://www.efas.eu/en>
- Joint Research Centre - European Commission. (2023b). *Global Flood Awareness System*. <https://www.globalfloods.eu/>
- Joint Research Centre - European Commission. (2023c). *Open Source Lisflood*.
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23(10), 4323-4331. <https://doi.org/10.5194/hess-23-4323-2019>
- Koch, J., & Schneider, R. (2022). Long short-term memory networks enhance rainfall-runoff modelling at the national scale of Denmark. *GEUS Bulletin*, 49, 1-7. <https://doi.org/10.34194/geusb.v49.8292>
- Kraft, B., Jung, M., Körner, M., Koirala, S., & Reichstein, M. (2022). Towards hybrid modeling of the global hydrological cycle. *Hydrology and Earth System Sciences*, 26(6), 1579-1614. <https://doi.org/10.5194/hess-26-1579-2022>
- Kratzert, F., Gauch, M., Nearing, G., & Klotz, D. (2022). NeuralHydrology — A Python library for Deep Learning research in hydrology. *Journal of Open Source Software*, 7(71), 4050. <https://doi.org/10.21105/joss.04050>
- 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., Herrnegger, M., & Hochreiter, S. (2018). *A glimpse into the Unobserved: Runoff simulation for ungauged catchments with LSTMs*.
- Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S., & Nearing, G. S. (2019). Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning. *Water Resources Research*, 55(12), 11344-11354. <https://doi.org/10.1029/2019WR026065>
- Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., & Matias, Y. (2023). Caravan - A global community dataset for large-sample hydrology. *Scientific Data*, 10(1). <https://doi.org/10.1038/s41597-023-01975-w>
- Lees, T., Reece, S., Kratzert, F., Klotz, D., Gauch, M., De Bruijn, J., Kumar Sahu, R., Greve, P., Slater, L., & Dadson, S. J. (2022). Hydrological concept formation inside long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 26(12), 3079-3101. <https://doi.org/10.5194/hess-26-3079-2022>
- Lehner, B., Messenger, M. L., Korver, M. C., & Linke, S. (2022). Global hydro-environmental lake characteristics at high spatial resolution. *Scientific Data*, 9(1). <https://doi.org/10.1038/s41597-022-01425-z>
- Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., & Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Nature Scientific Data*, 6(283). <https://doi.org/10.1038/s41597-019-0300-6>
- Liu, Y., Qin, H., Zhang, Z., Yao, L., Wang, Y., Li, J., Liu, G., & Zhou, J. (2019). Deriving reservoir operation rule based on Bayesian deep learning method considering multiple uncertainties. *Journal of Hydrology*, 579. <https://doi.org/10.1016/j.jhydrol.2019.124207>
- Mao, G., Wang, M., Liu, J., Wang, Z., Wang, K., Meng, Y., Zhong, R., Wang, H., & Li, Y. (2021). Comprehensive comparison of artificial neural networks and long short-term memory networks for rainfall-runoff simulation. *Physics and Chemistry of the Earth*, 123. <https://doi.org/10.1016/j.pce.2021.103026>

- Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting Through Conceptual Models Part I-a Discussion of Principles*. *Journal of Hydrology*, 10, 282-290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Nevo, S., Morin, E., Gerzi Rosenthal, A., Metzger, A., Barshai, C., Weitzner, D., Voloshin, D., Kratzert, F., Elidan, G., Dror, G., Begelman, G., Nearing, G., Shalev, G., Noga, H., Shavitt, I., Yuklea, L., Royz, M., Giladi, N., Peled Levi, N., ... Matias, Y. (2022). Flood forecasting with machine learning models in an operational framework. *Hydrology and Earth System Sciences*, 26(15), 4013-4032. <https://doi.org/10.5194/hess-26-4013-2022>
- Salamon, P., Grimaldi, S., Disperati, J., & Prudhomme, C. (2023). LISFLOOD static and parameter maps for GloFAS.
- Slater, L. J., Arnal, L., Boucher, M.-A., Chang, A. Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R. L., Wood, A., & Zappa, M. (2023). Hybrid forecasting: blending climate predictions with AI models. *Hydrology and Earth System Sciences*, 27(9), 1865-1889. <https://doi.org/10.5194/hess-27-1865-2023>
- Steyaert, J. C., Condon, L. E., W.D. Turner, S., & Voisin, N. (2022). ResOpsUS, a dataset of historical reservoir operations in the contiguous United States. *Scientific Data*, 9(1). <https://doi.org/10.1038/s41597-022-01134-7>
- Thiemig, V., Gomes, G. N., Skøien, J. O., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J. P., Kolbe, C., Pichon, D., Schweim, C., & Salamon, P. (2022). EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe. *Earth System Science Data*, 14(7), 3249-3272. <https://doi.org/10.5194/essd-14-3249-2022>
- Tsai, W. P., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., Liu, J., & Shen, C. (2021). From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-26107-z>
- Turner, S. W. D., Steyaert, J. C., Condon, L., & Voisin, N. (2021). Water storage and release policies for all large reservoirs of conterminous United States. *Journal of Hydrology*, 603. <https://doi.org/10.1016/j.jhydrol.2021.126843>
- van der Knijff, J. M., Younis, J., & de Roo, A. P. J. (2010). LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2), 189-212. <https://doi.org/10.1080/13658810802549154>
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11), 5844-5853. <https://doi.org/10.1002/2017GL072874>
- Yokoo, K., Ishida, K., Ercan, A., Tu, T., Nagasato, T., Kiyama, M., & Amagasaki, M. (2022). Capabilities of deep learning models on learning physical relationships: Case of rainfall-runoff modeling with LSTM. *Science of the Total Environment*, 802. <https://doi.org/10.1016/j.scitotenv.2021.149876>