

Redes neuronales

1.1. Modelo Denso

En primer lugar se ha considerado un modelo que consta de dos capas densas (Fig. 1.1).

Layer (type)	Output Shape	Param #
flatten_3 (Flatten)	(None, 9)	0
dense_31 (Dense)	(None, 200)	2000
dense_32 (Dense)	(None, 200)	40200
dense_33 (Dense)	(None, 76)	15276
Total params: 57,476		
Trainable params: 57,476		
Non-trainable params: 0		

Figura 1.1: Modelo Denso

Este modelo ha sido entrenado sobre una matriz que contiene las series temporales de precipitación, temperatura máxima y temperatura mínima para cada subcuenca y que posee la siguiente estructura:

$$X_{comp} = \begin{bmatrix} p_{1,1} & T_{1,1}^{max} & T_{1,1}^{min} & p_{1,2} & T_{1,2}^{max} & T_{1,2}^{min} & \dots & p_{1,nc} & T_{1,nc}^{max} & T_{1,nc}^{min} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p_{nt,1} & T_{nt,1}^{max} & T_{nt,1}^{min} & p_{nt,2} & T_{nt,2}^{max} & T_{nt,2}^{min} & \dots & p_{nt,n} & T_{nt,n}^{max} & T_{nt,n}^{min} \end{bmatrix}$$

Donde n_t es el número de pasos temporales y nc es el número de subcuencas. Los valores a ajustar son los caudales naturales simulados con el modelo hidrológico LEM y se encuentran almacenados en una matriz Y estructurada de la siguiente manera:

$$Y_{comp} = \begin{bmatrix} q_{1,1} & q_{1,2} & q_{1,2} \dots & q_{1,nc} \\ \dots & \dots & \dots & \dots \\ q_{nt,1} & q_{nt,2} & q_{nt,2} \dots & q_{nt,nc} \end{bmatrix}$$

Model: "sequential_13"

Layer (type)	Output Shape	Param #
lstm_10 (LSTM)	(None, 200)	168000
dense_29 (Dense)	(None, 200)	40200
dense_30 (Dense)	(None, 76)	15276

Total params: 223,476
 Trainable params: 223,476
 Non-trainable params: 0

Figura 1.2: Modelo LSTM 1

Si consideramos que la cuenca está compuesta por 76 subcuenas, y un rango temporal que va desde el 2001 hasta el 2020 con un paso temporal mensual, entonces las dimensiones de X_{comp} e Y_{comp} son 229x228 y 229x76, respectivamente. De estos datos se han tomado los primeros 160 como conjunto de entrenamiento y los últimos 69 como conjunto de test.

1.2. Modelo LSTM1

En segundo lugar se ha considerado un modelo que contiene una capa compuesta de una red recurrente y una segunda capa densa (Fig. 1.2) Se han utilizado dos maneras diferentes de entrenar este modelo, por un lado se ha entrenado sobre toda la cuenca, de la misma manera que el modelo denso, es decir utilizando las matrices X_{comp} e Y_{comp} , y por otro lado se ha entrenado el modelo individualmente cuenca por cuenca, para lo cual se han utilizado matrices X_{ind} e Y_{ind} con las siguientes formas:

$$X_{indv,id} = \begin{bmatrix} p_{1,id} & T_{1,id}^{max} & T_{1,id}^{min} \\ \dots & \dots & \dots \\ p_{nt,id} & T_{nt,id}^{max} & T_{nt,id}^{min} \end{bmatrix}$$

$$Y_{indv,id} = \begin{bmatrix} q_{1,id} \\ \dots \\ q_{nt,id} \end{bmatrix}$$

Donde id es el número de la subcuenca.

Model: "sequential_15"

Layer (type)	Output Shape	Param #
lstm_11 (LSTM)	(None, 1, 100)	41600
lstm_12 (LSTM)	(None, 100)	80400
dense_34 (Dense)	(None, 100)	10100
dense_35 (Dense)	(None, 1)	101

=====
Total params: 132,201
Trainable params: 132,201
Non-trainable params: 0
=====

Figura 1.3: Modelo LSTM 2

En este último caso, las matrices poseen las dimensiones 229x3 y 229x1, respectivamente. La división entre los conjuntos de entrenamiento y test se ha hecho de la misma manera que en el apartado anterior.

1.3. Modelo LSTM2

Por último se ha considerado un tercer modelo que consta de dos capas LSTM y una capa densa (Fig. 1.3) que se ha entrenado de manera secuencial utilizando únicamente los valores simulados de los caudales en la matriz $Y_{indv,id}$. En este enfoque la red utiliza los valores de los caudales en tiempos anteriores q_{t-1} , q_{t-2} , q_{t-n} para predecir el valor actual q_t . La variable n también llamada "look back" se puede elegir arbitrariamente y se puede ajustar haciendo una busca con el método de cross validation.

1.4. Resultados

En la figura 1.5 se muestran a modo de ejemplo los resultados obtenidos para los caudales de tres puntos diferentes de la cuenca, la sub-cuenca 1, la sub-cuenca 20 y la 73. En las figuras se muestran los resultados de los caudales simulados con el modelo hidrológico y los obtenidos con los modelos Denso y LSTM1 en el conjunto de test. Se puede apreciar que agregar una capa LSTM mejora ligeramente los resultados ya que se incrementa el valor del coeficiente Nash-Sutcliffe (NSE) para el modelo. Esto era de esperar ya que la red LSTM tiene en consideración la dependencia temporal de las diferentes entradas mientras que el modelo denso considera las entradas de los distintos pasos de tiempo como independientes.

En la figura 1.4 se muestran las funciones de autocorrelación correspondientes a cada una de la subcuencas. Se puede observar que ambos modelos

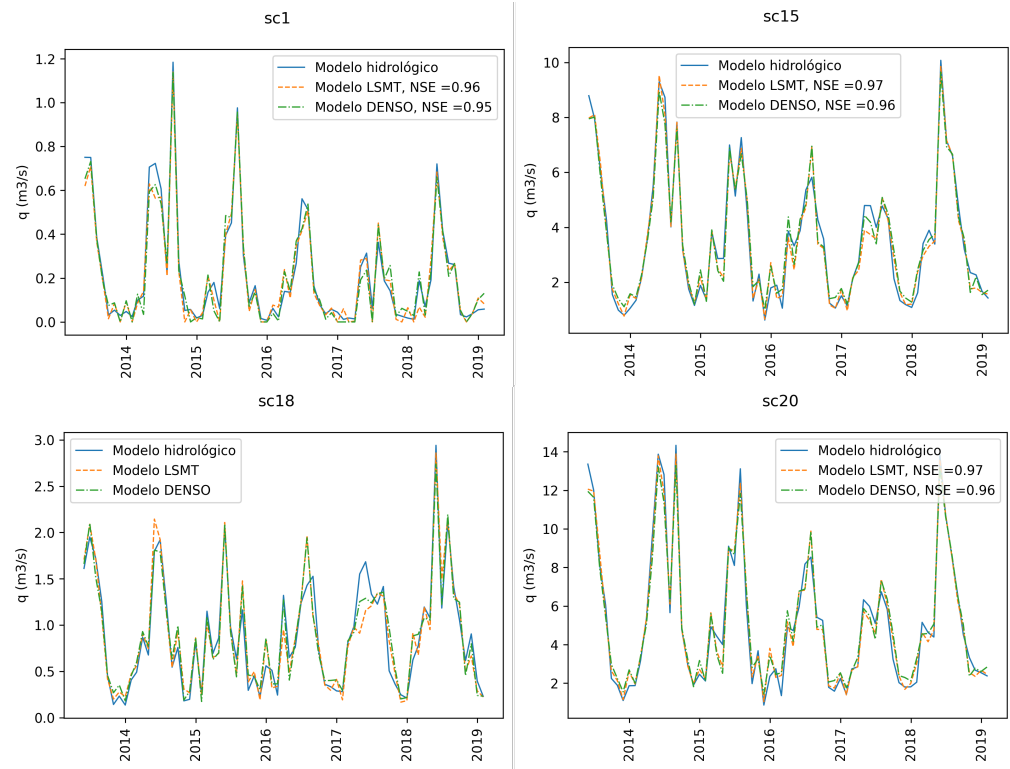


Figura 1.4: Resultados del Modelo LSTM 2.

son capaces de captar correctamente la estacionalidad de los datos.

En la figura 1.6 se muestran los resultados obtenidos con el modelo LSTM3 en todo el rango de datos para la subcuenca numero 73. En este enfoque, se han predicho los datos de manera secuencial, para lo cual sólo se ha utilizado la información de las series temporales de los caudales simulados en el conjunto de entrenamiento. Si bien este modelo predice correctamente la tendencia de los datos, falla al predecir los valores de los caudales.

1.4.1. Validación de los modelos

Para validar los modelos y hacer un análisis un poco más profundo de cómo es su performance en los diferentes puntos de la cuenca, se ha utilizado el coeficiente de Nash-Sutcliffe (NSE). En la figura 1.7 se muestran los valores obtenidos de NSE en todas las subcuencas para el modelos LSTM1. La curva punteada muestra los valores obtenidos al entrenar el modelo en cada subcuenca de forma individual, la curva dashed es la obtenida entrenando el modelo para toda la cuenca en conjunto considerando la matriz X_{comp} que posee 288 características. Se puede observar que para este modelo, el valor de

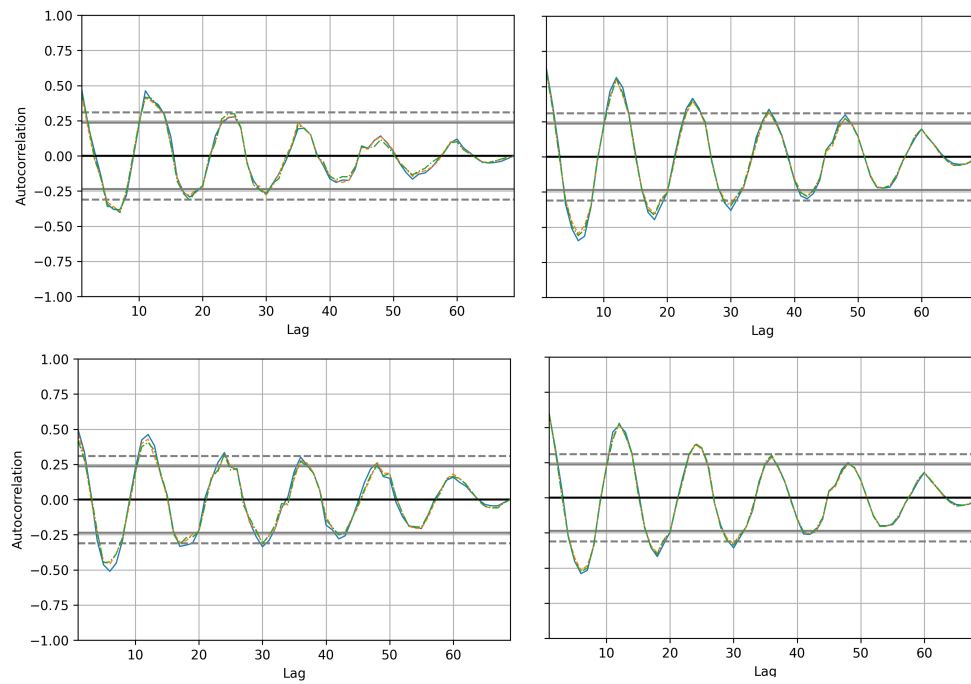


Figura 1.5: Funciones de autocorrelación.

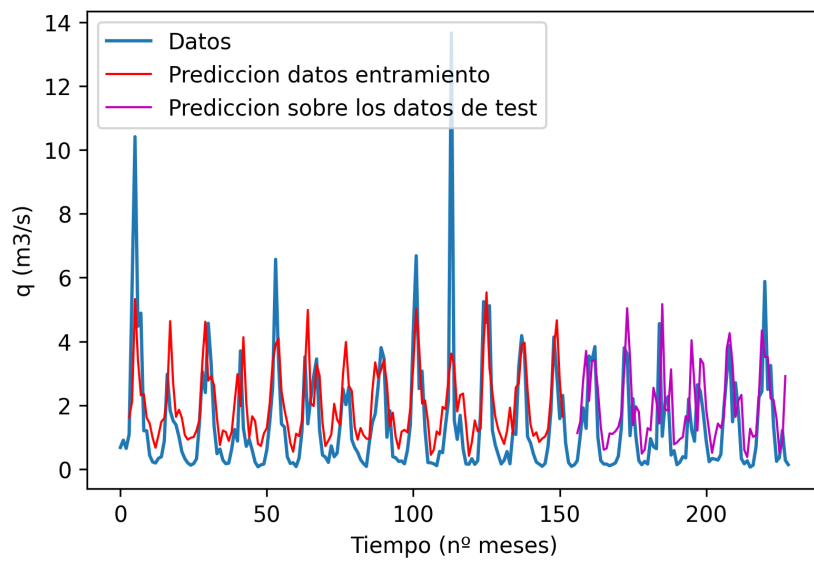


Figura 1.6: Resultados del modelo LSTM 3.

NSE baja considerablemente en muchos puntos, esto se debe a que al haber

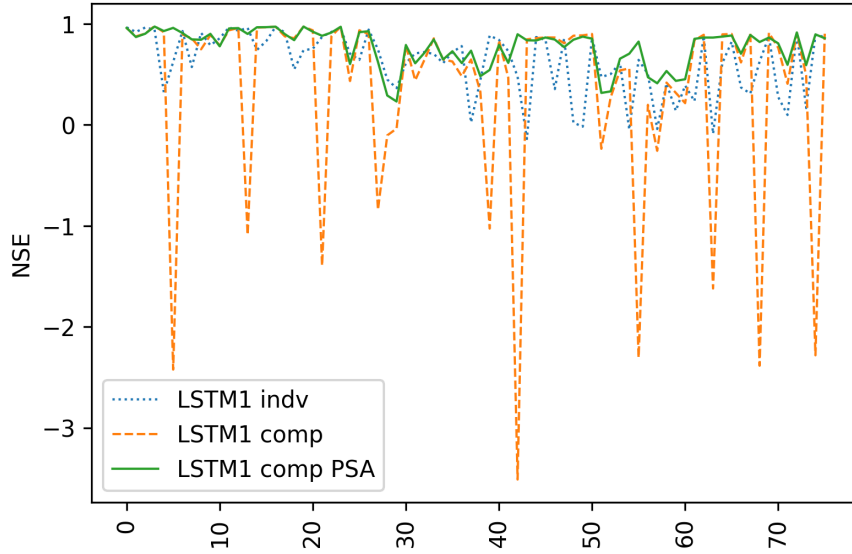
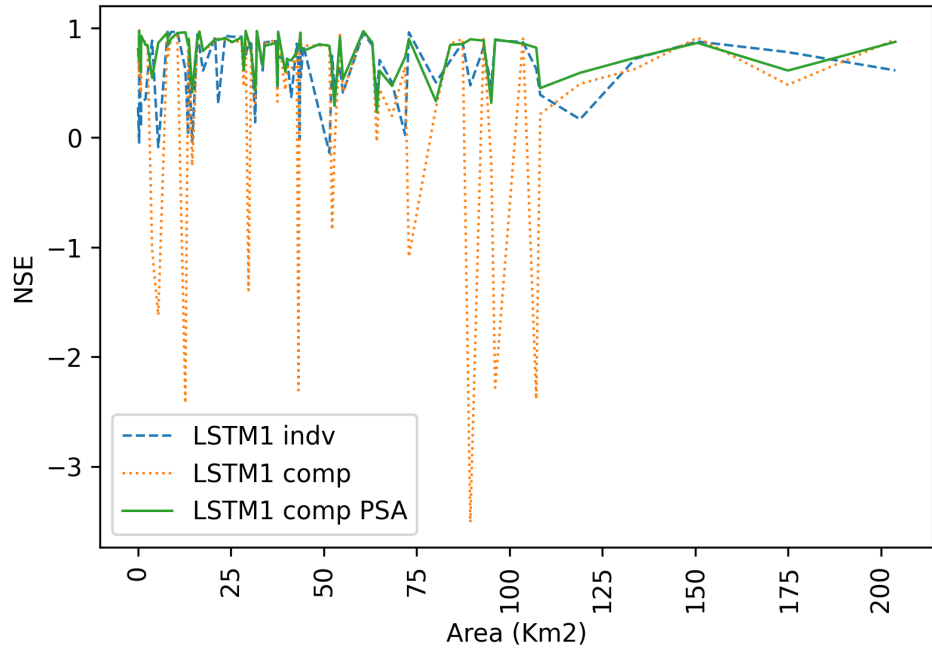


Figura 1.7: Modelo LSTM 2

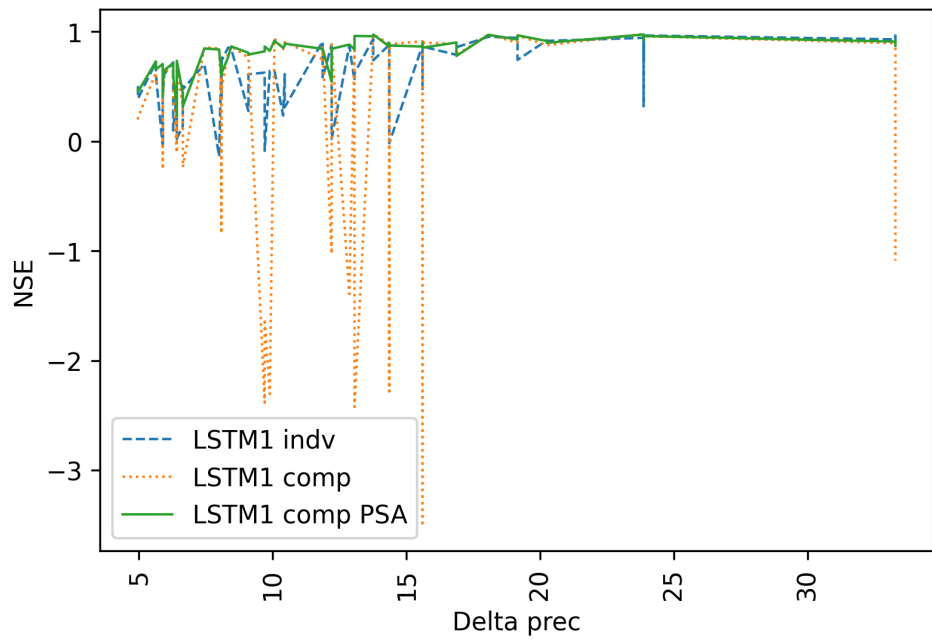
tantas características de entrada se produce un sobre ajuste de los datos. Esto se puede corregir haciendo una reducción del espacio considerando las componente principales (curva sólida). En este análisis el número de características se reduce a 9 componentes principales que explican el 96 % de la varianza de los datos.

Con el fin de determinar si hay otros factores que provocan el sobre ajuste de los datos, en la figura 1.8 se muestran los valores de NSE en función del área de las cuencas (panel izquierdo) y de la variación máxima de precipitación (panel derecho). Se puede observar que los modelos tienden a fallar para cuencas con áreas pequeñas ($A < 125 \text{ km}^2$) y cuando la serie de precipitación no muestra grandes variaciones ($\Delta P < 20 \text{ mm/d}$). Esto se debe a que en este caso es difícil para los modelos poder distinguir los verdaderos patrones de precipitación del ruido, una alternativa para mejorar estos resultados sería considerar una capa de autoencoder al inicio de los modelos que permitan filtrar el ruido en los datos.

Este efecto es aún más visible cuando entrenamos los modelos de manera local, si comparamos la curva continua con la dashed, podemos ver que en la mayoría de los casos, los resultados obtenidos entrenando el modelo en toda la cuenca son mejores que cuando lo hacemos de manera individual. Esto se debe a que en el primer caso estamos considerando los patrones de precipitación en diferentes puntos de la cuenca, y aunque localmente la variación de la precipitación sea pequeña, la red aprende a completar la información faltante utilizando la información de toda la cuenca. En el



(a)



(b)

Figura 1.8: Coeficientes de Nash-Sutcliffe en función del área (1.8a) y la variación de presión (1.8b).

segundo caso sólo consideramos los patrones locales en una cuenca dada y si la información otorgada por la serie de precipitación no es buena (la variación es pequeña y no se reconoce un patrón claro) la red no es capaz de completar la información necesaria con información proveniente de otros puntos de la cuenca.

1.5. Resultados finales

Con el fin de determinar si los resultados obtenidos utilizando redes neuronales pueden ser utilizados en un problema real de gestión de suministros en una cuenca, se ha alimentado el programa Modsim con los caudales simulados por el modelo hidrológico LEM y los caudales obtenidos por el modelo LSTM1 en el conjunto de test. Se han considerado los diferentes usos del agua por demandas humanas, riego, industriales, etc. Los resultados para la salida de la cuenca, se muestran en la figura 1.9. Se puede apreciar que los resultados son prácticamente iguales.

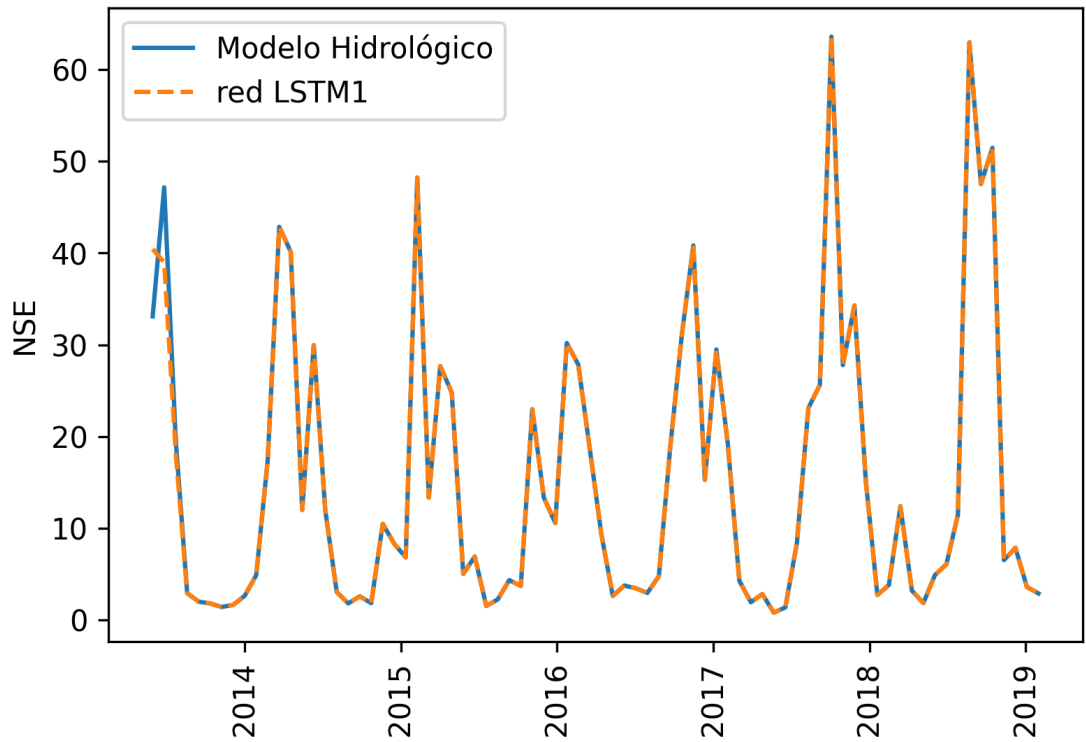


Figura 1.9: Resultados finales obtenidos utilizando el software de gestión de recursos Modsim, teniendo en cuenta los usos del agua