Informe

Carmen Calvo Olivera

2 de junio de 2020

Resumen

En este documento se recogen los resultados obtenidos de la aplicación de diferentes técnicas de aprendizaje automático a nuestro conjunto de datos, cuyo objetivo es conseguir un modelo capaz de... calcular en tiempo real el índice de verosimilitud para una predicción meteorológica a partir de la obtencion del error cuadrático medio entre la precipitacón obtenida por el modelo WRF (Weather Research and Forecasting) y nuestro modelo.

Para ello, se ha llevado a cabo la creación de un conjunto de datos a partir de predicciones obtenidas por el modelo WRF.

1. Introducción

En la actualidad, ... \mathbb{R} $\hat{\mathbb{R}}$

2. Dataset

La preparación de los datos supone una parte esencial a la hora de trabajar con aprendizaje automático. En esta sección de describre brevemente tanto la obtención como el tratamiento (o preprocesamiento) de las predicciones a partir de las cuales se obtienen los conjuntos de datos utilizados para el entrenamiento y validación de los distintos clasificadores.

En primer lugar, cabe destacar uso de dos grandes conjuntos de datos, uno para train y el otro para validation.

- train_dataset: este conjunto de datos abarca desde febrero (a la expera de añadir enero) de 2015 hasta diciembre de 2015.
- validation_dataset: que abarca desde enero de 2016 hasta diciembre de 2016.

2.1. Obtención

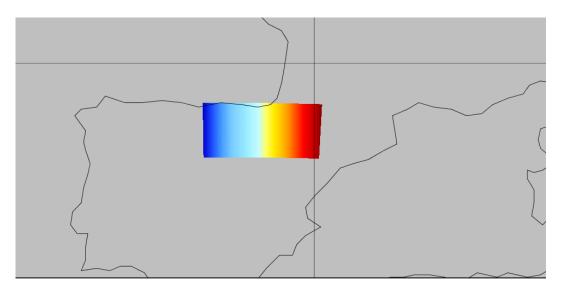
Los datos de las predicciones se obtinen a partir del modelo WRF, un modelo meteorológico numérico de mesoescala no hidrostático, utilizado para finalidades tanto de predicciones operativas en tiempo real como para investigación atmosférica.

Los datos de inicialización del modelo WRF provienen de análisis y predicciones de otros NWP cuyo formato debe ser GRIB1 o GRIB2. En nuestro caso, todos los ficheros fuente han sifo obtenido de Research Data Archive del NCAR (Centro Nacional de Investigación Atmosférica).

Posteriormete se llevan a cabo todos los pasos incluidos dentro del flujo de procesamiento y, a continuación, comienza la simulación meterorológica. Tras todo esto, obtenemos una serie de ficheros del tipo " $wrfout_d0X_yyyy-mm-dd_hh:mm:ss$ "donde X es el número del dominio, y la cadena yyyy-mm-dd_hh:mm:ss representa la fecha y hora de la primera salida guardada en el fichero. Todos estos ficheros son obtenidos en formato NetCDF, un formato de archivo destinado a almacenar datos científicos multidimensionales (variables) como la temperatura, la humedad o la presión.

En nuestro caso, se hará uso de los ficheros horarios para el dominio 2 representado en la Figura 1, y definido previamente en los ficheros de configuración, para 24h, es decir, para el día 1 de enero de 2016 se hace uso de 24 ficheros horarios que van desde "wrfout_d02_2016-01-02_01:00:00" hasta "wrfout_d02_2016-02-03_00:00:00".

Figura 1. Representación del dominio de predicción correspondiente a la zona del Ebro



2.2. Procesamiento

Tras la obtención de las predicciones, se llevan a cabo una serie de pasos para el tratamiento de los datos y finalmente la obtención de nuestro dataset:

■ Filtración de características: el primer paso llevado a cabo es la filtración de las variables que obtenemos de las predicciones del WRF y obtener así una seleccion de variables detalladas en la tabla Tabla 1. Mediante el uso de la libreria de python wrf-python te hace un tratamiento de los datos para, a partir de una prediccion horaria (por qué horaria (?)), obtener un único fichero .nc diario con

toda la información que necesitamos. A continuación se muestra la cabecera de unos de los ficheros filtrados:

```
[ccalvo@frontend1 nc] $ ncdump -h 2016-04-12.nc
netcdf \setminus 2016-04-12 {
dimensions:
        south_north = 78;
        west_east = 123;
        time = 24;
variables:
        float XLAT(time, south_north, west_east);
        float XLONG(time, south_north, west_east);
        float HGT(time, south_north, west_east);
        float RAINC(time, south_north, west_east);
        float RAINNC(time, south_north, west_east);
        string DATE(time);
        float TIMESTAMP(time);
        float QVAPOR_500(time, south_north, west_east)
        float QVAPOR_700(time, south_north, west_east)
        float QVAPOR_850(time, south_north, west_east)
        float QCLOUD_500(time, south_north, west_east)
        float QCLOUD_700(time, south_north, west_east)
        float QCLOUD_850(time, south_north, west_east);
        float QRAIN_500(time, south_north, west_east);
        float QRAIN_700(time, south_north, west_east);
        float QRAIN_850(time, south_north, west_east);
        float QICE_500(time, south_north, west_east);
        float QICE_700(time, south_north, west_east);
        float QICE_850(time, south_north, west_east);
        float QSNOW_500(time, south_north, west_east);
        float QSNOW_700(time, south_north, west_east);
        float QSNOW_850(time, south_north, west_east);
        float QGRAUP_500(time, south_north, west_east);
        float QGRAUP_700(time, south_north, west_east);
        float QGRAUP_850(time, south_north, west_east);
        float T<sub>-</sub>500(time, south_north, west_east);
        float T_700(time, south_north, west_east);
```

Tabla 1. Variables utilizadas para la creación de un dataset

Variable	Descripción
DATE	Fecha de la predicción
TIMESTAMP	

Variable	Descripción
XLAT	
XLONG	
HGT	
RAINC	Precipitación convectiva
RAINNC	Precipitación no convectiva
T_500hPa	Temperatura a diferentes presiones
$T_{-}700hPa$	-
T_850hPa	
$QVAPOR_{-}500$	Razón de mezcla
QVAPOR_700	
QVAPOR_850	
$QCLOUD_{-}500$	
$QCLOUD_{-700}$	
$QCLOUD_{-850}$	
QRAIN_500	
$QRAIN_{-}700$	
QRAIN_850	
$QICE_500$	
$QICE_{-700}$	
$QICE_{-}850$	
$QSNOW_{-}500$	
$QSNOW_{-700}$	
$QSNOW_850$	
$QGRAUP_500$	
QGRAUP_700	
QGRAUP_850	

• Creación de los csv: el siguiente paso es la creación de un dataset, para el cual los ficheros NetCDF se convierten en formato CSV y posterioremnte se añaden algunas variables (o etiquetados de los datos). En la tabla Tabla 2 se recogen las variables añadidas a nuestros ficheros.

Tabla 2. Variables añadidas a los datasets

Variable	Descripción
	Precipitación acumulada de la
PRECIPITACION_WRF	predicción del WRF (RAINC +
	RAINNC)
	Precipitación acumulada real
PRECIPITACION	(obtenida a partir pluviómetros de la
	CHE)
	continúa en la siguiente página

Variable	Descripción
	Variable binaria para la predicción del
LLUVIAWRF	WRF (0 \rightarrow No precipitación y 1 \rightarrow
	precipitación)
LLUVIA	Variable binaria para la CHE (0 \rightarrow No
LLUVIA	precipitación y $1 \rightarrow \text{precipitación}$)
RANGO_WRF	Rango para la predicción del WRF ([0
RANGO_WRF	- 14]*)
RANGO	Rango para la CHE ($[0 - 14]^*$)

^{*} Rangos (mm): 0.1,1.,1.5,2.5,5.,10.,15.,20.,25.,30.,40.,50.,80. Rangos (representación): [0-14]

■ Creación de un único csv: como último paso, y previo al entrenamiento de todos los modelos, se crean dos ficheros .csv con los días que deseamos incluir en nuestro dataset de *train* o de *validation*.

3. Resultados

A continuación, se detallan los resultados obtenidos, los cuales se han organizado de la siguiente manera. Todos ellos cuentan con la tasa de acierto obtenida en el conjunto de datos destinado a la validación tanto en el train como en el test y posteriormente con el error cuadrático medio.

El error cuadrático medio o mean square error (MSE) corresponde al promedio de los errores al cuadrado, es decir, la diferencia entre el valor real y el valor estimado. Se corresponde con:

$$MSE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{sample}-1} (y_i - \hat{y}_i)^2$$

Donde \hat{y}_i corresponde al valor predicho para cada *i*-muestra e y_i corresponde al valor real.

- Binaria: en primer lugar se muestran los resultados obtenido para la predicción de la variable binaria con 3 escenarios distintos.
 - Todas las variables incluidas en el CSV a excepción de aquellas consideradas como etiquetas. Ver Tabla 2. Resultados: Tabla 3
 - Todas las variables del punto anterior eliminando RAINC y RAINNC. Resultados: Tabla 4
 - Únicamente con las coordenadas y las variables RAINC y RAINNC. Resultados: Tabla 5

- Rango: en este caso se realiza la predicción con los rangos definidos con los 3 escenarios anteriores.
 - Todas las variables incluidas en el CSV a excepción de aquellas consideradas como etiquetas. Ver Tabla 2. Resultados: Tabla 6
 - Todas las variables del punto anterior eliminando RAINC y RAINNC. Resultados: Tabla 7
 - Únicamente con las coordenadas y las variables RAINC y RAINNC. Resultados: Tabla 8

El error cuadrático medio correspondiente a la relación entre el valor predicho por el WRF y el valor real obtenido de los pluviómetros sería el siguiente:

Error cuadrático medio (WRF - CHE)= 0.34860156086592764

Tras este cáculo, analizamos las predicciones obtenidas de muestros clasificadores respecto de la verdad terreno y calculamos nuevamente el MSE.

Tabla 3. Binaria. Todas características

			MSE
Variable	Train	\mathbf{Test}	(model -
			CHE)
MPL	0.736403	0.756870	0.243130
LogisticRegression	0.728701	0.738289	0.261711
QDA	0.703161	0.731465	0.268535
NeuralNetwork	0.743430	0.731454	0.268546
LDA	0.740250	0.731177	0.268823
OVR	0.688272	0.727229	0.272771
AdaBoost	0.770249	0.721126	0.278874
RandomForest	0.696645	0.712761	0.287239
DecisionTree	0.773973	0.711559	0.288441
NaiveBayes	0.732275	0.705945	0.294055
KNeighbors	0.892685	0.645533	0.354467
SGDClassifier	0.393367	0.358522	0.641478

Tabla 4. Binaria. Todas características sin RAIN

			Error
Variable	Train	\mathbf{Test}	cuadrático
			\mathbf{medio}
SGDClassifier	0.686392	0.715450	0.284550
	0.00002	0.110400	0.20

Variable	Train	Test	Error cuadrático
			\mathbf{medio}
LDA	0.734469	0.711946	0.288054
QDA	0.688710	0.704254	0.295746
LogisticRegression	0.709877	0.698761	0.301239
DecisionTree	0.742524	0.698344	0.301656
NeuralNetwork	0.715451	0.694780	0.305220
RandomForest	0.669924	0.690709	0.309291
OVR	0.668298	0.690013	0.309987
MPL	0.718232	0.675204	0.324796
AdaBoost	0.730053	0.672329	0.327671
NaiveBayes	0.709254	0.630211	0.369789
KNeighbors	0.862658	0.617016	0.382984

Tabla 5. Binaria. Solo RAIN

Variable	Train	Test	Error cuadrático medio
MPL	0.741252	0.753171	0.246829
NeuralNetwork	0.744667	0.750569	0.249431
LogisticRegression	0.714578	0.749263	0.250737
NaiveBayes	0.712043	0.748209	0.251791
LDA	0.703919	0.741426	0.258574
DecisionTree	0.764922	0.719928	0.280072
AdaBoost	0.762199	0.717789	0.282211
RandomForest	0.753338	0.715067	0.284933
QDA	0.667406	0.687748	11.073562
KNeighbors	0.813913	0.678051	0.321949
OVR	0.426253	0.407584	0.592416
SGDClassifier	0.333523	0.313059	0.686941

 ${\bf Tabla~6.} \quad {\sf Rango.~Todas~features}$

Variable	Train	Test	Error cuadrático medio
OVR	0.669560	0.689178	9.145259
MPL	0.669085	0.688222	9.102521
NeuralNetwork	0.669087	0.688164	9.079000

Variable	Train	Test	Error cuadrático
			\mathbf{medio}
RandomForest	0.667468	0.687758	11.070078
LogisticRegression	0.668037	0.684587	8.619534
SGDClassifier	0.664522	0.677533	8.401055
LDA	0.664978	0.675906	9.561050
AdaBoost	0.666912	0.655973	9.656310
DecisionTree	0.680278	0.651697	8.835721
NaiveBayes	0.651177	0.620479	8.792356
KNeighbors	0.850602	0.541724	10.216860
QDA	0.414040	0.394458	64.326082

Tabla 7. Rango. Todas features sin RAIN

			Error
Variable	Train	Test	cuadrático
			${f medio}$
NeuralNetwork	0.667406	0.687749	11.072704
MPL	0.667406	0.687748	11.073562
RandomForest	0.667415	0.687747	11.073196
LogisticRegression	0.665773	0.685467	10.581673
AdaBoost	0.667699	0.677845	11.002875
LDA	0.662489	0.675684	10.928471
DecisionTree	0.673827	0.662918	10.344793
SGDClassifier	0.654264	0.653064	10.036488
NaiveBayes	0.636050	0.564641	10.302585
KNeighbors	0.809269	0.535653	11.650298
QDA	0.303942	0.293877	83.531754
OVR	0.262102	0.229779	11.982337

Tabla 8. Rango. Solo RAIN

Variable	Train	Test	Error cuadrático medio
RandomForest	0.667406	0.687748	11.073562
QDA	0.667406	0.687748	11.073562
OVR	0.668867	0.687748	9.380225
NeuralNetwork	0.668598	0.687300	9.158468
MPL	0.668488	0.686898	9.082758

Variable	Train	Test	Error cuadrático medio
LogisticRegression	0.667769	0.683597	8.929119
LDA	0.664579	0.681965	10.134818
AdaBoost	0.668623	0.677344	9.771984
DecisionTree	0.670085	0.674760	9.498254
NaiveBayes	0.665110	0.673433	8.316886
KNeighbors	0.786454	0.592601	9.774854
SGDClassifier	0.148744	0.135874	10.248119

Por último, se ha llevado a cabo el entrenamiento de los tres mejores clasificadores para predecir el valor binario con todas las características sin incluir las variables de precipitación RAINC y RAINNC (un total de 24 variables), con diferente número de características para así intentar obtener mejores resultados. La Tabla 9 recoge los resultados obtenidos correspondientes al clasificador LDA, en la Tabla 10 se recogen los resultados para el clasificados QDA y en la Tabla 11 para el clasificador SGD. A mayores, se han incluido los resultados obtenido para el clasificador NeuralNetwork en la Tabla 12.

Tabla 9. Resultados para k características. LDA

\overline{k}	Features	Train	Test	\mathbf{ECM}
	XLAT XLONG HGT			
	$T_500hPa\ T_700hPa$			
	$T_850hPa~QVAPOR_500$	0.734522		
	QVAPOR_700		0.715925	0.284075
	$QVAPOR_{-850}$			
20	QCLOUD_{-700}			
20	QCLOUD_850 QRAIN_850			
	$QICE_{500} \ QICE_{700}$			
	$QICE_{850} QSNOW_{500}$			
	$QSNOW_{-700} QSNOW_{-850}$			
	QGRAUP_700			
	QGRAUP_850			

\boldsymbol{k}	Features	Train	\mathbf{Test}	\mathbf{ECM}
23	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_500 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.734466	0.712566	0.287434
21	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.734249	0.712048	0.287952

\boldsymbol{k}	Features	Train	Test	\mathbf{ECM}
22	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_850	0.734458	0.711965	0.288035
19	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.734515	0.708498	0.291502
18	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.734944	0.706961	0.293039
3	T_500hPa T_700hPa T_850hPa	0.714001	0.705687	0.294313

\boldsymbol{k}	Features	Train	Test	\mathbf{ECM}
14	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.717245	0.687752	0.312248
10	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.716688	0.687748	0.312252
15	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.717292	0.687748	0.312252
17	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.735030	0.595058	0.404942
16	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.735131	0.535027	0.464973
4	XLAT T_500hPa T_700hPa T_850hPa	0.714624	0.312252	0.687748

\boldsymbol{k}	Features	Train	\mathbf{Test}	\mathbf{ECM}
5	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850	0.716302	0.312252	0.687748
6	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QSNOW_700	0.716045	0.312252	0.687748
7	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.716373	0.312252	0.687748
8	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.716763	0.312252	0.687748
9	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.716719	0.312252	0.687748
11	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.716780	0.312252	0.687748
12	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850	0.717129	0.312252	0.687748
13	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.717213	0.312252	0.687748

 ${f Tabla~10.}~~{\sf Resultados~para}~k~{\sf características.}~{\sf QDA}$

\overline{k}	Features	Train	Test	ECM
21	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.687764	0.704529	0.295471
22	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.688060	0.704364	0.295636
16	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.685077	0.703971	0.296029

$oldsymbol{k}$	Features	Train	Test	\mathbf{ECM}
15	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.684032	0.703930	0.296070
20	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.686165	0.702544	0.297456
23	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_500 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.688333	0.700021	0.299979
12	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850	0.687094	0.699652	0.300348

\boldsymbol{k}	Features	Train	Test	ECM
17	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.685408	0.699503	0.300497
3	T_500hPa T_700hPa T_850hPa	0.713155	0.694569	0.305431
14	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.683711	0.688878	0.311122
13	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.685100	0.688147	0.311853
10	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.687190	0.687748	0.312252
11	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.687207	0.687748	0.312252

$oldsymbol{k}$	Features	Train	Test	ECM
18	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.685491	0.681203	0.318797
19	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.685936	0.680043	0.319957
4	XLAT T_500hPa T_700hPa T_850hPa	0.718118	0.312252	0.687748
5	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850	0.699664	0.312252	0.687748
6	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QSNOW_700	0.689058	0.312252	0.687748
7	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.690120	0.312252	0.687748
8	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.690843	0.312252	0.687748
9	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.687227	0.312252	0.687748

 ${f Tabla\ 11.}$ Resultados para k características. Neural Network

$-\frac{1}{k}$	Features	Train	Test	ECM
17	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.675428	0.688741	0.311259
23	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_500 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.667474	0.687889	0.312111
14	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.677734	0.687883	0.312117
10	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.667451	0.687748	0.312252

k	Features	Train	Test	ECM
15	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.699792	0.687748	0.312252
18	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.667388	0.687748	0.312252
16	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.332594	0.681350	0.318650
20	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.702025	0.663193	0.336807

\boldsymbol{k}	Features	Train	Test	\mathbf{ECM}
22	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.707176	0.659079	0.340921
19	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.682827	0.654114	0.345886
3	T_500hPa T_700hPa T_850hPa	0.624041	0.546839	0.453161
21	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.572248	0.508894	0.491106

${m k}$	Features	Train	Test	\mathbf{ECM}
4	XLAT T_500hPa T_700hPa T_850hPa	0.667484	0.312252	0.687748
5	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850	0.507912	0.312252	0.687748
6	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QSNOW_700	0.604810	0.312252	0.687748
7	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.668499	0.312252	0.687748
8	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.509600	0.312252	0.687748
9	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.481599	0.312252	0.687748
11	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.676183	0.312252	0.687748
12	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850	0.332670	0.312252	0.687748
13	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.667747	0.312252	0.687748

Tabla 12. Resultados para k características. Neural Network

$-\frac{1}{k}$	Features	Train	Test	ECM
15	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.680742	0.710796	0.289204
14	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QSNOW_700 QSNOW_850 QGRAUP_850	0.709281	0.705374	0.294626
16	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.715923	0.702667	0.297333
19	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.712400	0.699214	0.300786

\boldsymbol{k}	Features	Train	Test	\mathbf{ECM}
23	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_500 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.718361	0.694276	0.305724
20	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.717065	0.693742	0.306258
18	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_850	0.717560	0.692513	0.307487

$oldsymbol{k}$	Features	Train	Test	ECM
22	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_500 QGRAUP_500 QGRAUP_500	0.716623	0.690795	0.309205
8	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.718454	0.687748	0.312252
9	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.721118	0.687748	0.312252
10	XLAT XLONG T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.723849	0.687748	0.312252
11	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QSNOW_700 QSNOW_850	0.697577	0.687748	0.312252

\boldsymbol{k}	Features	Train	Test	\mathbf{ECM}
17	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_700 QSNOW_850 QGRAUP_850	0.720185	0.682017	0.317983
3	T_500hPa T_700hPa T_850hPa	0.717968	0.680073	0.319927
21	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QVAPOR_500 QVAPOR_700 QVAPOR_850 QCLOUD_700 QCLOUD_850 QRAIN_700 QRAIN_850 QICE_500 QICE_700 QICE_850 QSNOW_500 QSNOW_700 QSNOW_850 QGRAUP_700 QGRAUP_850	0.719092	0.666388	0.333612
4	XLAT T_500hPa T_700hPa T_850hPa	0.712840	0.312252	0.687748
5	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850	0.713077	0.312252	0.687748
6	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QSNOW_700	0.702289	0.312252	0.687748
7	XLAT T_500hPa T_700hPa T_850hPa QCLOUD_850 QRAIN_850 QSNOW_700	0.719495	0.312252	0.687748
12	XLAT XLONG HGT T_500hPa T_700hPa T_850hPa QCLOUD_700 QCLOUD_850 QRAIN_850 QICE_700 QSNOW_700 QSNOW_850	0.719892	0.312252	0.687748

${m k}$	Features	Train	\mathbf{Test}	\mathbf{ECM}
	XLAT XLONG HGT			
	$T_500hPa\ T_700hPa$			
	$T_850hPa\ QCLOUD_700$			
13	QCLOUD_850 QRAIN_850	0.682436	0.312252	0.687748
	$QICE_{700} QSNOW_{700}$			
	$QSNOW_{-}850$			
	QGRAUP_850			

4. Trabajo futuro

Como trabajo inmediato posterior, se pretende complementar los resultados con la siguiente información:

- Incluir enero de 2015 en el dataset de entrenamiento.
- Realizar predicciones con lo mm.
- Hacer pruebas para obtener el mejor score para distintos KBest.

5. Referencias