

Time series model selection from a
bayesian practical approach

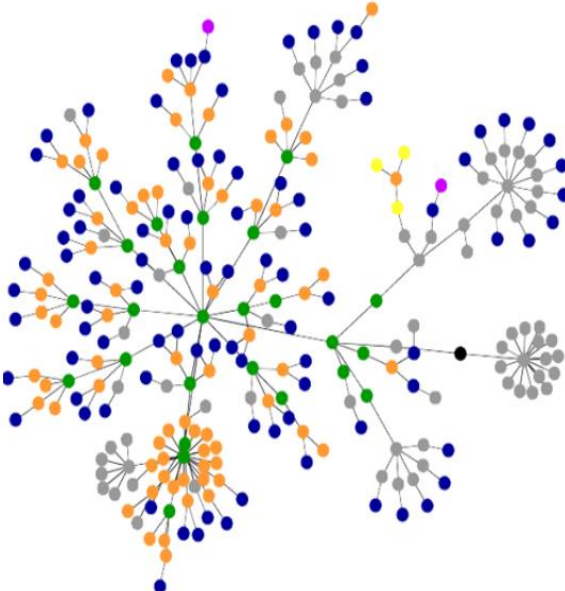
Explicación de la aplicación Shiny expuesta en la comunicación tipo B

← → ↻ 127.0.0.1:7464 🔍 ☆

Time series

Introducción
Settings
Description
Ajuste y predicción
Bayesian Network
Source Code
Bookmark...

Red bayesiana para identificar el modelo que a priori mejor predice una serie temporal



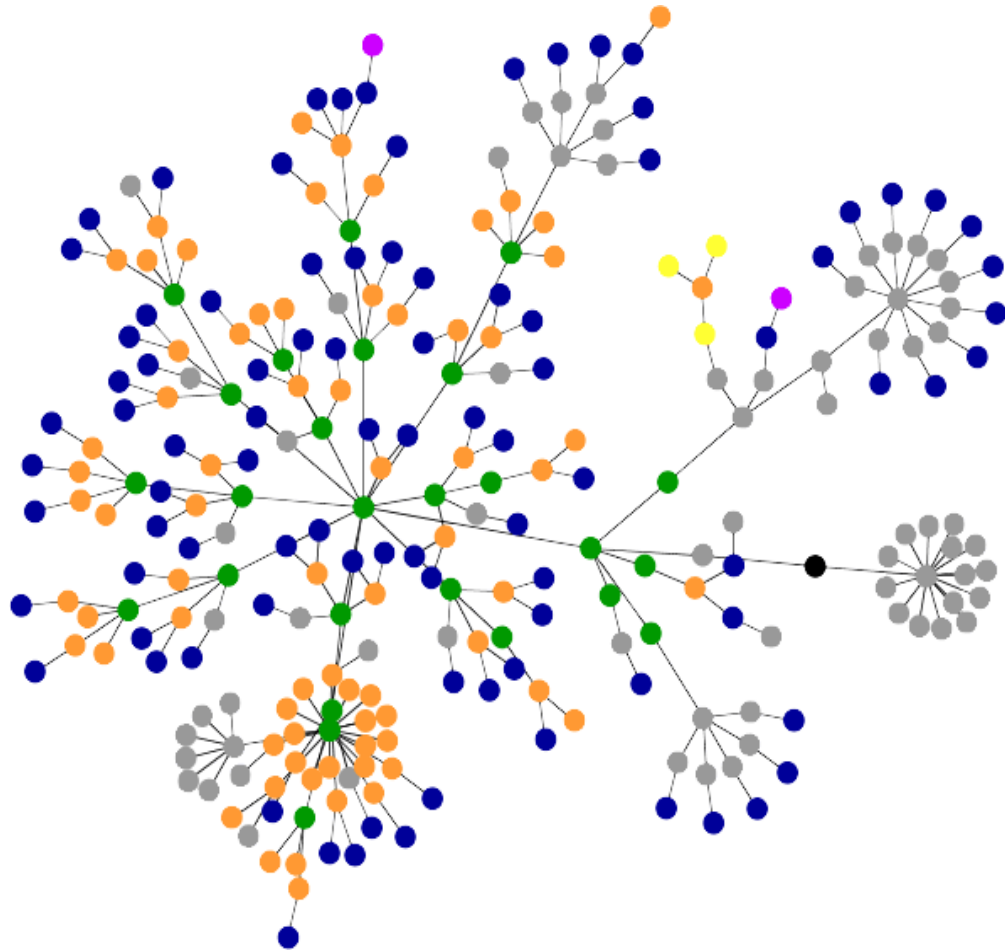
El objetivo de este trabajo es:
Dada una serie temporal con sus características iniciales, identificar el modelo que menor error de predicción *a priori* comete.

Los modelos de predicción que se han empleado son:

- ETS
- BSTS
- Prophet
- ARIMA

Pasos a seguir

Red bayesiana para identificar el modelo que a priori mejor predice una serie temporal



El objetivo de este trabajo es:

Dada una serie temporal con sus características iniciales, identificar el modelo que menor error de predicción *a priori* comete.

Los modelos de predicción que se han empleado son:

ETS
BSTS
Prophet
ARIMA

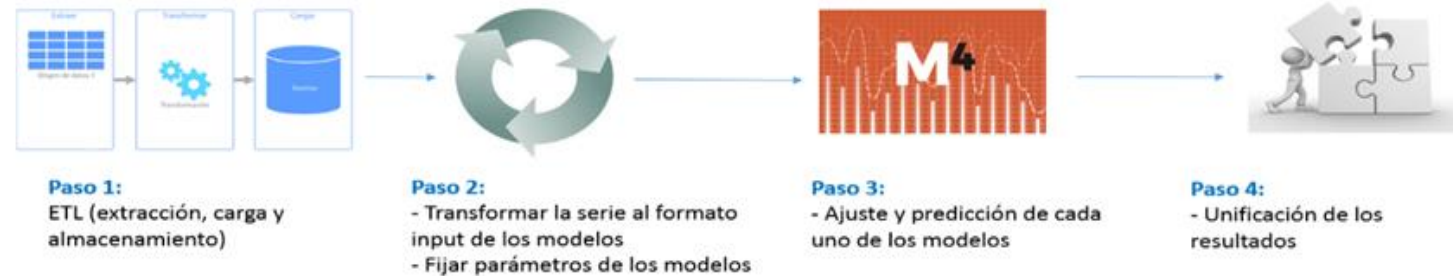
Para contrastar los modelos se ha empleado una red bayesiana

1. Tabla de eventos

Tabla de eventos			Granularidad mensual	Series temporales		
Fecha	ID Cliente	Saldo		Fecha	ID Cliente	Saldo
2017/05/06	A	X€	→	2017/05	A	X€
2017/05/06	B	T€		2017/06	A	X€
2017/08/13	A	Y€		2017/07	A	X€
2017/08/13	A	Z€		2017/08	A	$(Z+Y)/2€$
2017/08/21	B	H€		2017/05	B	T€
				2017/06	B	T€
				2017/07	B	T€
				2017/08	B	H€

Paso 1: Elegir la granularidad
Paso 2: Cálculo de agregados por granularidad
Paso 3: Búsqueda del mínimo y máximo de las series
Paso 4: Relleno de valores missing

2. Descrip,ajust,predic



3. Comparación resultados de los modelos

Análisis
descriptivo de
los resultados

Categorización
MAPE y
longitud

Creación de
una red
bayesiana

Descripción del dataset

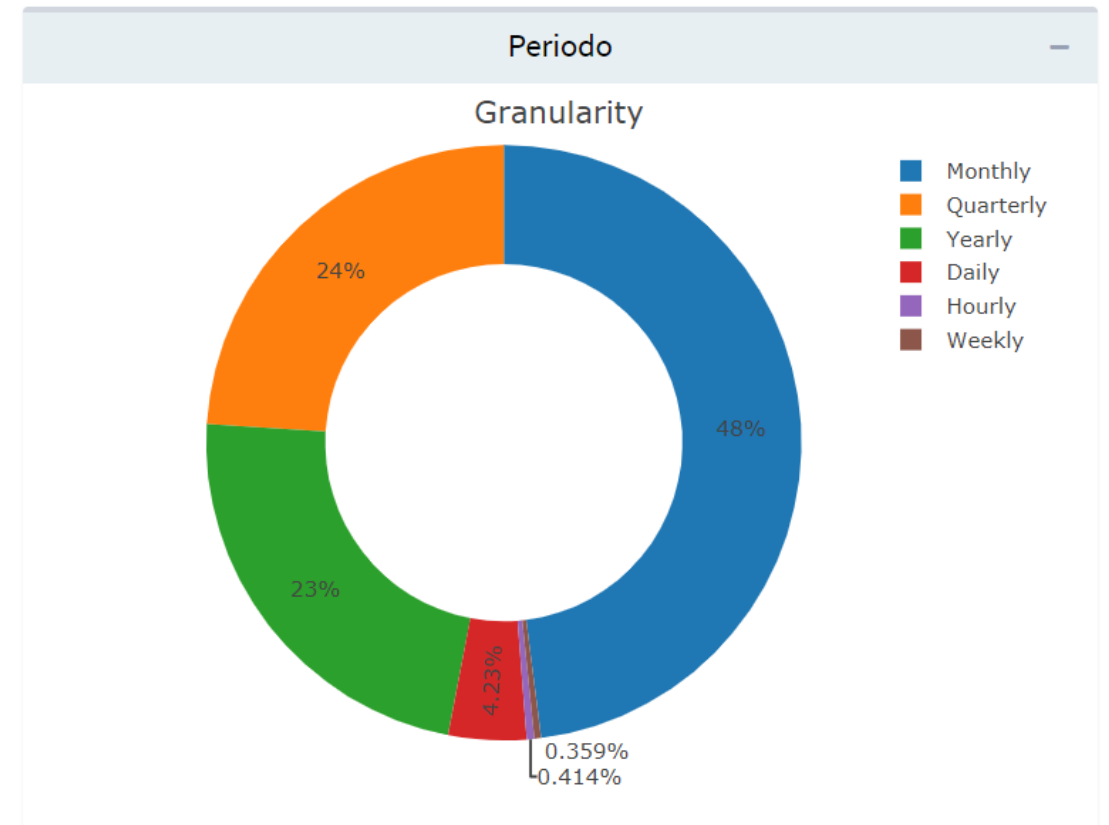
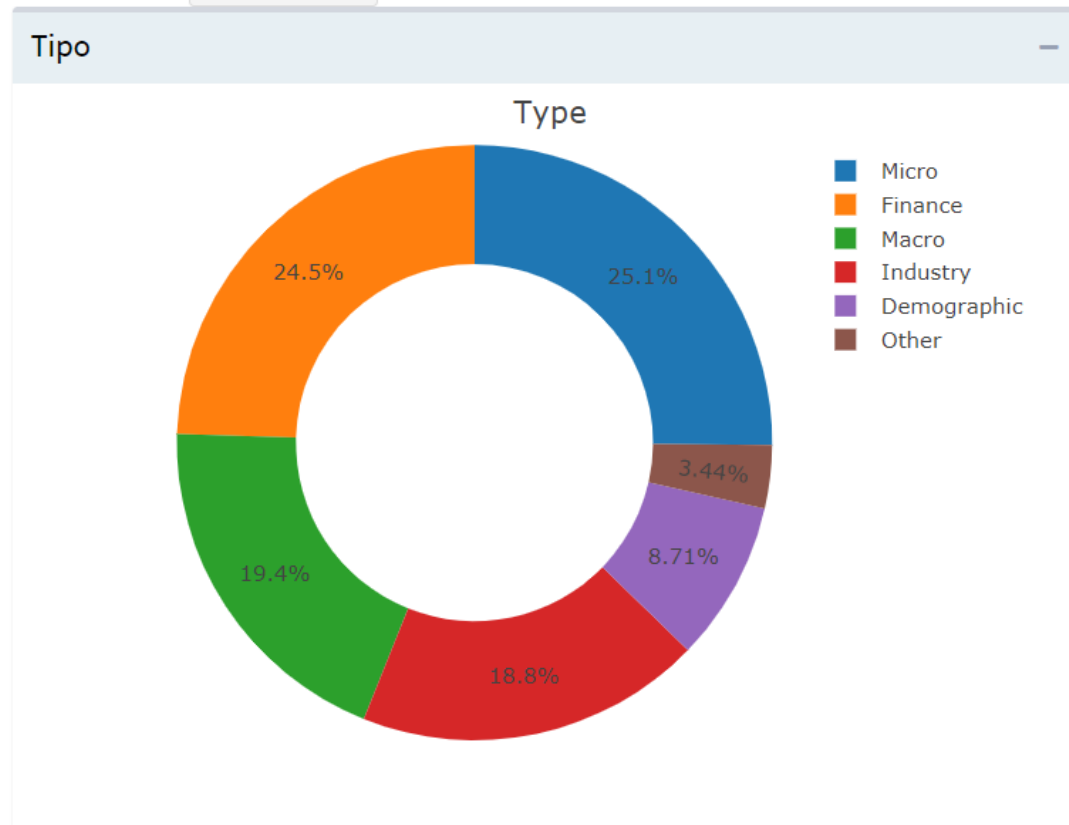
El dataset empleado tiene un total de 100.000 series temporales de diferentes tipos (demograficas, financieras, industriales, micro, macro y otras) y distintas granularidades (anuales, trimestrales, mensuales, semanales, diarias, horarias).

La muestra de test empleada es 6 periodos para las series anuales, 8 para las trimestrales, 18 para las mensuales, 13 para las semanales, 14 para las diarias y 48 para las horarias.

La muestra para predecir ha sido del total de series para los modelos ARIMA y ETS y la mitad de series temporales para los modelos BSTS y Prophet debido a su elevado tiempo de ejecucion.

Descargar dataset:

 M4 dataset



Número de series dentro de cada segmento

Frecuency table

Show 10 ▾ entries

Search:

	Demographic ↕	Finance ↕	Industry ↕	Macro ↕	Micro ↕	Other ↕	Total ↕
Yearly	1088	6519	3716	3903	6538	1236	23000
Quarterly	1858	5305	4637	5315	6020	865	24000
Monthly	5728	10987	10017	10016	10975	277	48000
Weekly	24	164	6	41	112	12	359
Daily	10	1559	422	127	1476	633	4227
Hourly	0	0	0	0	0	414	414
Total	8708	24534	18798	19402	25121	3437	100000

Showing 1 to 7 of 7 entries

Previous

1

Next

Caso particular: Análisis de dos series temporales

Tipo: microeconómicas

Periodo: mensual

Parametros

Coefficiente de la tendencia polinomial

1

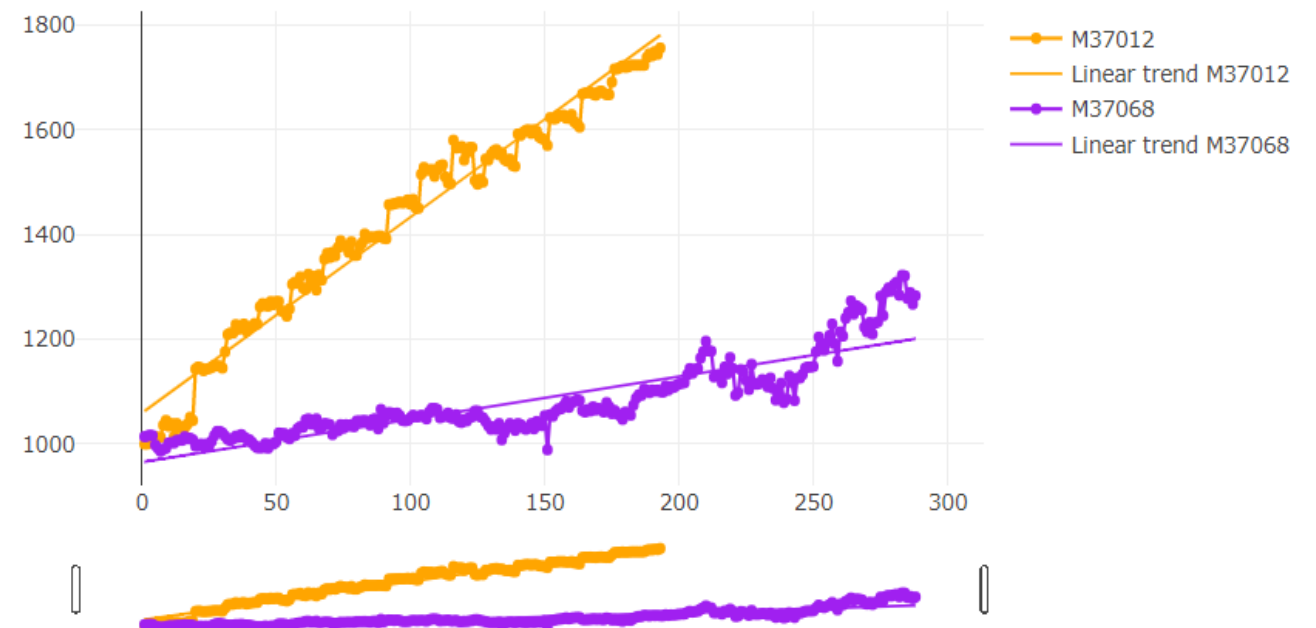
numero de diferencias

0

ACF y PACF: Selecciona una serie

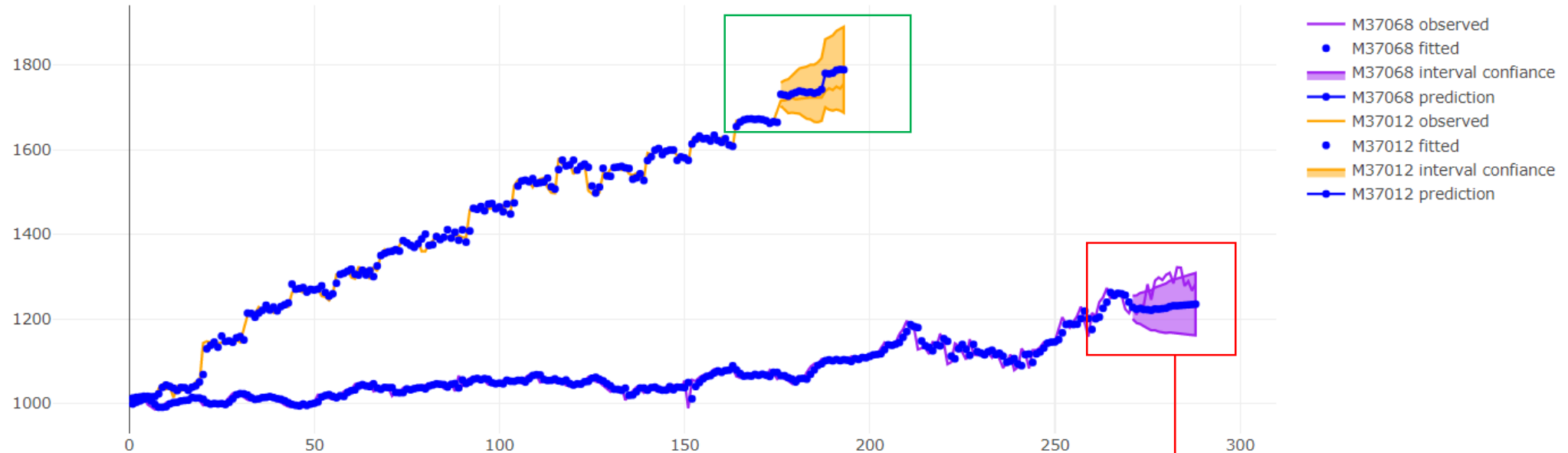
M37012

Series temporales seleccionadas



Ajuste y predicción del modelo ARIMA

Model ARIMA

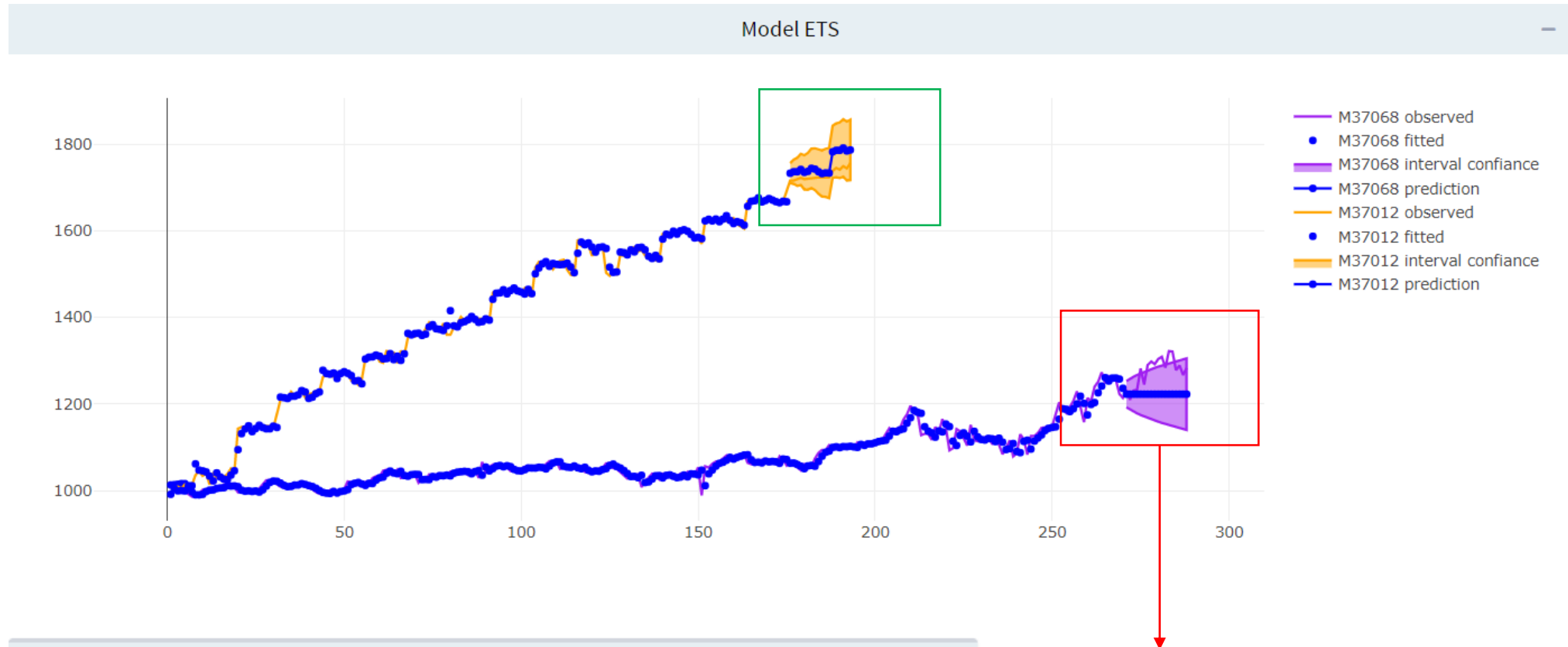


Accuracy test

	ME	RMSE	MAE	MPE	MAPE	ACF1
M37068	48.66	57.41	50.11	3.76	3.88	0.60
M37012	-21.81	25.20	21.81	-1.26	1.26	0.79

El intervalo de confianza no abarca la mayoría de las predicciones

Ajuste y predicción del modelo ETS

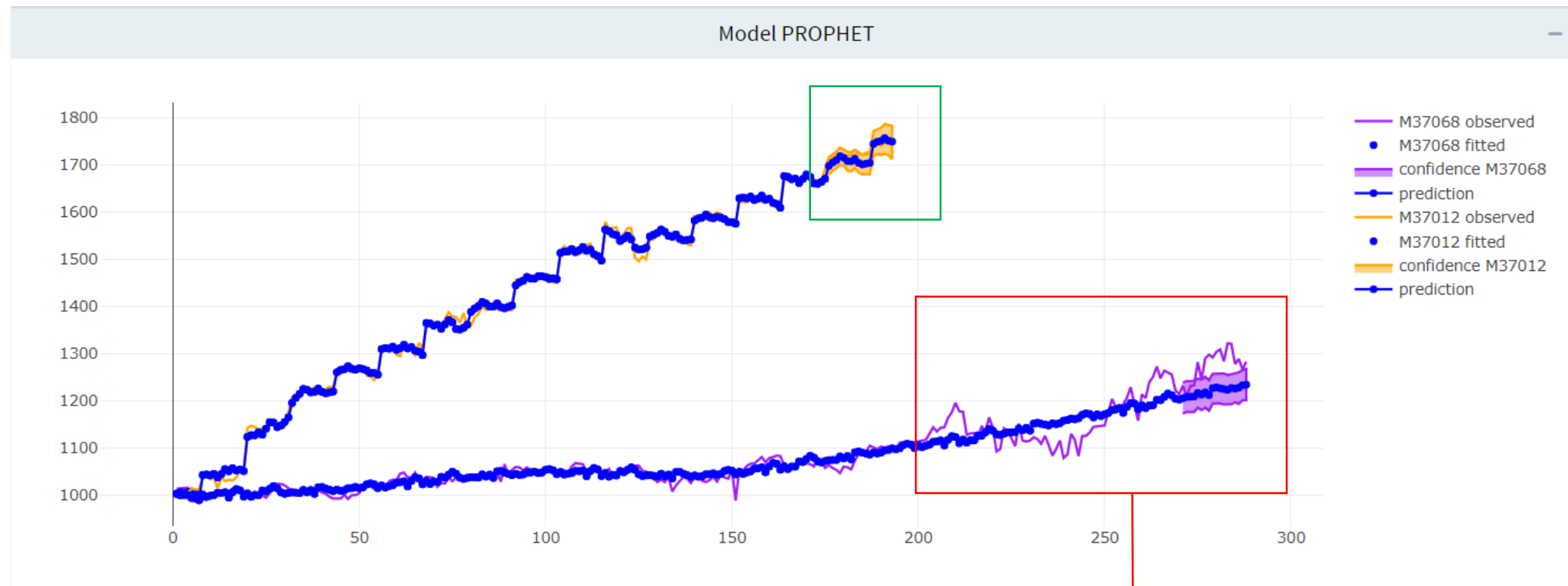


Accuracy test

	ME	RMSE	MAE	MPE	MAPE	MASE
M37068	53.54	62.39	54.94	4.13	4.25	0.63
M37012	-23.97	26.97	23.97	-1.38	1.38	0.72

El intervalo de confianza no abarca la mayoría de las predicciones

Ajuste y predicción del modelo Prophet

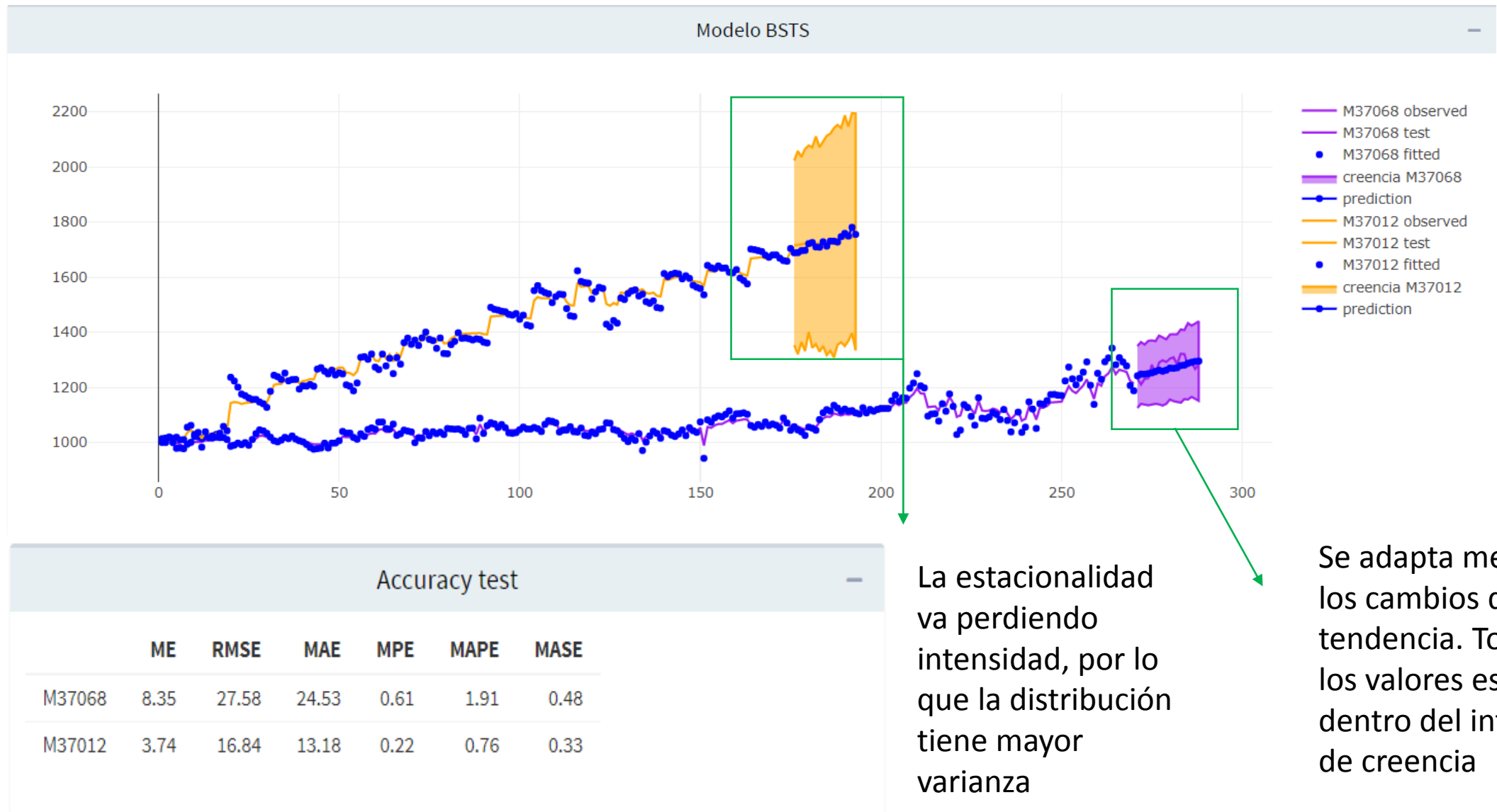


Accuracy test

	ME	RMSE	MAE	MPE	MAPE	MASE
M37068	55.41	61.50	55.41	4.29	4.29	0.53
M37012	7.29	12.78	11.37	0.42	0.66	0.67

No detecta los puntos de cambio de tendencia debido a que no se han podido introducir por ser un proceso automatizado de predicción de 100000 series

Ajuste y predicción del modelo BSTS



Red bayesiana discreta para contrastar los resultados

¿Por qué una red bayesiana?

- Las redes bayesianas buscan relaciones causales. En nuestro caso, se quiere identificar las relaciones entre las características iniciales de las series temporales y las medidas de error de la predicción con cada uno de los modelos
- Se puede introducir información a priori tanto en la estructura de la red como en las probabilidades iniciales

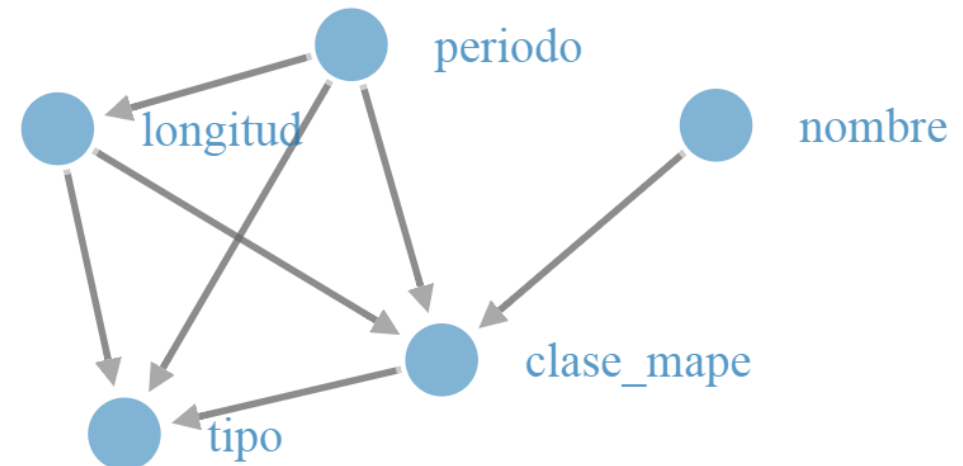
Score de la red

Score de la red:

Bayesian Information Criterion

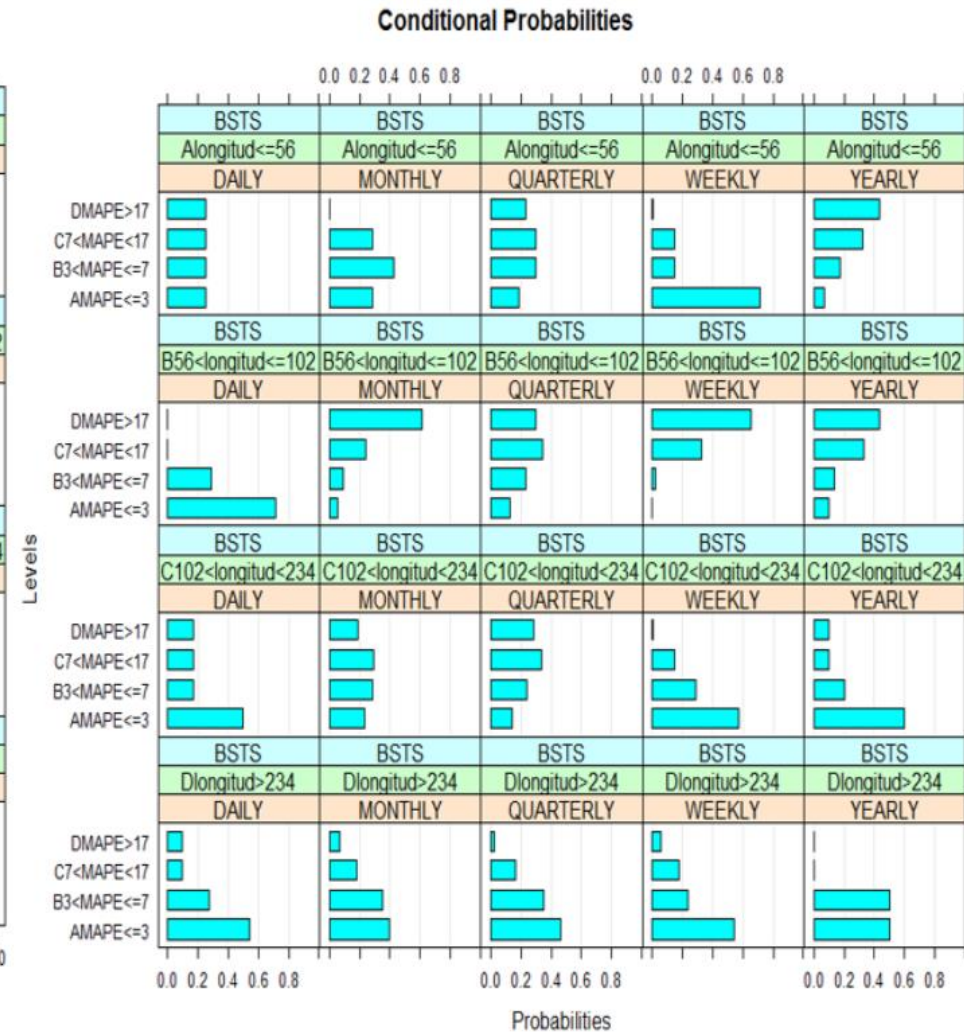
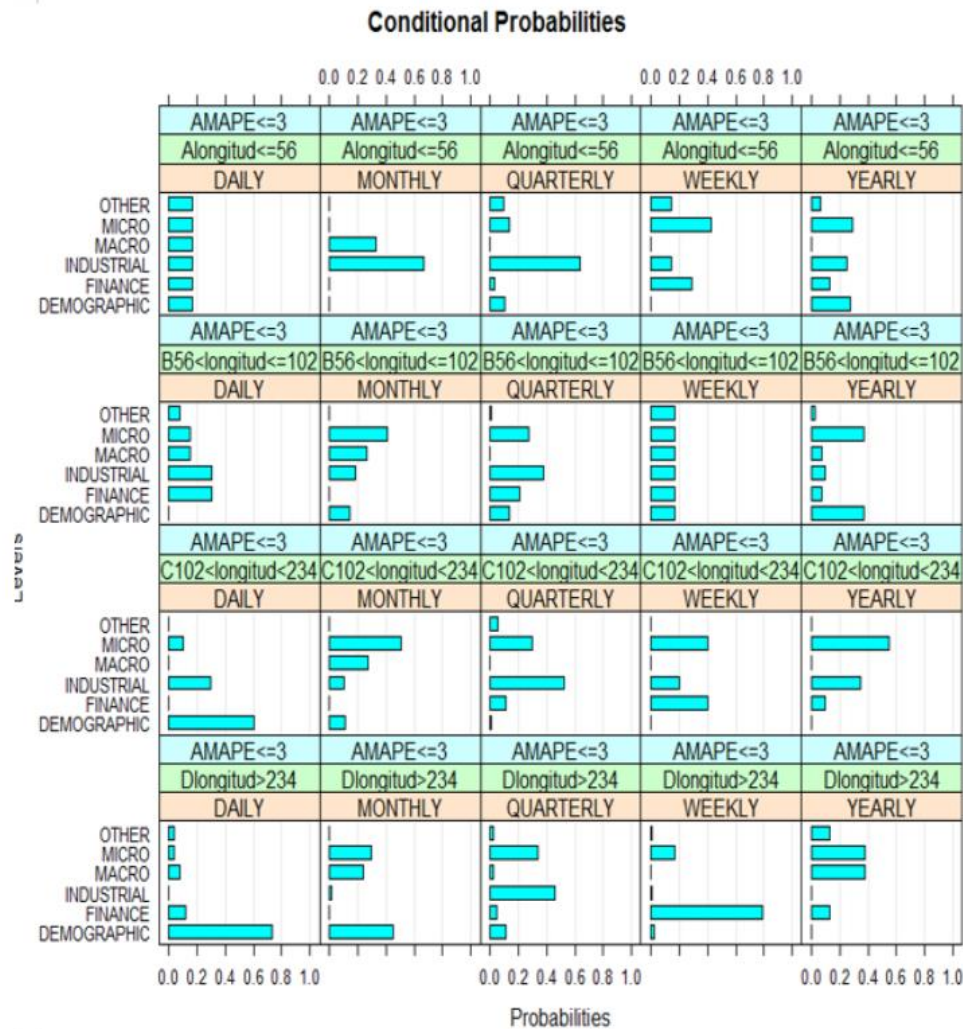
-564060.9

Red Bayesiana



Algunas probabilidades obtenidas en la red

Probabilidades condicionadas



Inferencia

Dado un conjunto de evidencias (categorías de variables que se conocen), se puede calcular la probabilidad de que ocurra un evento (en este caso el evento es que la variable MAPE tenga un error menor de 3 y que el modelo empleado sea ARIMA, BSTS, ETS o Prophet)



Conclusiones

Resultados obtenidos



<https://github.com/AnaBotey/Bayesian-network-to-compare-time-series-models>

Ana Arias Botey

- Los modelos ETS y ARIMA tienen un tiempo de ejecución considerablemente menor que los modelos Prophet y BSTS.
- El modelo Prophet es el que mejor actúa para series de tipo diario de longitudes medias y largas.
- Los modelos ARIMA, ETS y BSTS obtienen buenos resultados en los segmentos de longitudes largas.
- Las predicciones del modelo BSTS son las que mejores resultados generales obtienen.