

OPTIMIZACIÓN DE LAS ACTUACIONES DEL CUERPO DE BOMBEROS DE LA COMUNIDAD DE MADRID

ALEJANDRO RASERO Y JUNHAO GE

RESUMEN. En este documento se presenta el informe de los procedimientos realizados en el proyecto relativo a la asignatura de *Descubrimiento de conocimiento de Datos Complejos*. Se explicarán los modelos aplicados y los resultados obtenidos.

1. INTRODUCCIÓN

El objetivo de esta práctica es predecir y optimizar las actuaciones del **Cuerpo de Bomberos** de la comunidad de Madrid mediante herramientas de análisis y predicción de series temporales vistas en la asignatura [1]. Se usará como información complementaria el registro de las actuaciones del **cuerpo de policía** y el **SAMUR** para emplearlas como variables exógenas.

Todos los datos se han obtenido del *Portal de Datos Abiertos de la Comunidad de Madrid*, los cuales se han extraído en formato *.xlsx* desde el año 2017 hasta el año 2023. La implementación completa se puede encontrar en el repositorio de *Github*[2].

2. ANÁLISIS DE LOS DATOS

Como se ha mencionado, se han empleado 3 datases diferentes:

- Actuaciones del Cuerpo de Bomberos [3]
- Policía Municipal. Datos estadísticos actuaciones Policía Municipal [4]
- Activaciones del SAMUR-Protección Civil [5]

Tanto los datos relativos a los bomberos como los del SAMUR se obtuvieron en formato *.xlsx*, por lo que ha sido necesario un ligero preprocesamiento, tanto para cambiar la codificación del texto como para exportar los archivos en formatos *.csv*. Cada uno de los ficheros referencian a los datos de un año en concreto, que será identificado por el nombre del mismo (p.e. "*ActuacionesBomberos_2017*") y el preprocesamiento realizado se encuentra en el notebook **CargaSamurBomberos.ipynb**.

Los datos del cuerpo de policía están separados por meses en lugar de años, por lo que ha sido necesario juntarlos antes de realizar el análisis. También se han detectado ligeros cambios en el formato de los datos entre los años, tanto por la inclusión de nuevas columnas, como por cambios en los nombres en ciertas variables. Por ello, se ha realizado un pequeño preprocesamiento para estandarizar los formatos de los dataframes hasta obtener un dataset completo y filtrado. Véase el notebook **limpieza_datos_policia.ipynb**.

Antes de analizar cada conjunto de forma independiente, se ha hecho una visualización general sobre las distribuciones de cada cuerpo a lo largo de los años [imagen 1].

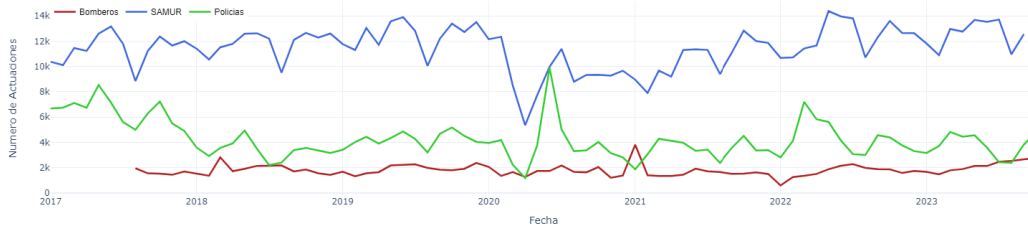


FIGURA 1. Intervenciones de los tres cuerpos en el periodo 2017-2023

Se puede apreciar que existe una mayor correlación entre la policía y el SAMUR, compartiendo en muchas ocasiones la tendencia de subidas y bajadas de incidencias a lo largo de los años. En cambio, la distribución del cuerpo de bomberos es más estable respecto a los demás, presentando poca variación salvo en ciertas excepciones que se analizarán más adelante.

Con este primer análisis se puede esperar que sea ciertamente difícil encontrar una correlación entre los bomberos y los otros dos cuerpos. Pero dado a que comparten ciertos periodos de aumentos y descensos, obviando cambios sutiles, es preferible investigar la posibilidad de usarlos como variables exógenas por si se pueden realizar mejores predicciones.

Cabe destacar que los datos del cuerpo de policía y del SAMUR alcanzan su mínimo histórico en enero de 2021 mientras que dicho mes es el máximo global de la serie que recoge las actuaciones de los bomberos. Curiosamente, este hecho no se corresponde con el patrón que se ha seguido hasta el momento. Este repunte repentino se debe a la borrasca Filomena que experimentó España en las primeras semanas de enero lo que multiplicó por 4 el número de intervenciones de los bomberos [6] [7], mientras que los otros cuerpos no se vieron realmente afectados, es más, se produjo el efecto contrario.

3. ANÁLISIS DEL CUERPO DE BOMBEROS

Visto la comparativa [*imagen 1*] es entendible suponer que los bomberos son aquellos con el menor número de intervenciones por la propia naturaleza de la profesión, considerando que a diferencia de estos, la policía y el SAMUR son cuerpos que intervienen en situaciones más comunes.

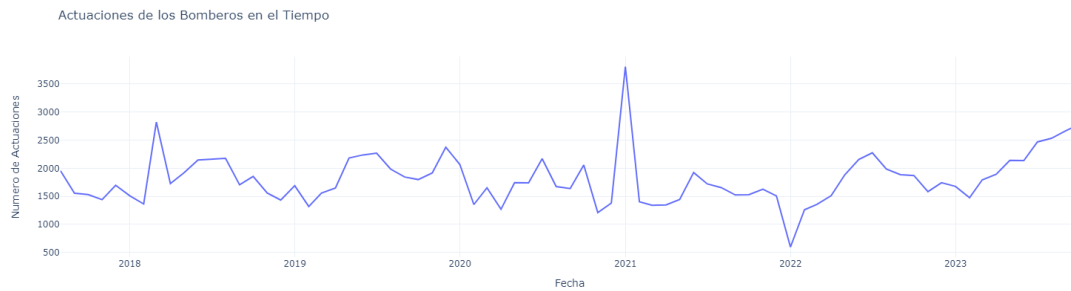


FIGURA 2. Actuaciones de los Bomberos en el periodo 2017-2023

Las actuaciones de los bomberos [*imagen 2*] generalmente aumentan en las épocas de verano y descienden en invierno. Aunque hay 3 casos anómalos que caben destacar:

1. **Marzo de 2018** presentó un repunte de incidencias que no se corresponde con la época del año. Según el informe de la Comunidad de Madrid [8] fue a causa de las grandes tormentas de dicho mes, junto a intervenciones por viento, lluvia, e incendios producidos en el sur.
2. **Enero de 2021** es el máximo histórico, que como se ha comentado es debido a la borrasca Filomena.
3. **Enero de 2022** presenta un mínimo histórico de intervenciones, que coincide con una reducción del número de intervenciones de la policía y el SAMUR, fue un mes tranquilo.

La época de COVID-19 produjo una subida en los últimos meses de 2019, justo antes de empezar el confinamiento. Pero en los meses posteriores de invierno no hubo un cambio significativo respecto a los mismos meses de otros años.

Cabe mencionar que el número de intervenciones ha empezado a aumentar considerablemente a partir de febrero de 2023. Un patrón que no se ha visto en los años anteriores, lo que dificultará las predicciones futuras.

3.1. Estudio de estacionalidad y Correlaciones. Para comprobar si la serie es estacionaria se ha realizado la prueba de *Dickey Fuller*[9], la cual otorgó un p_valor menor a 0.05, rechazando la hipótesis nula lo que sugiere que la serie es estacionaria. Por tanto, no será necesario integrar la serie.

En lo relativo al estudio de las correlaciones, se ha realizado la Autocorrelación simple (ACF)[*imagen 3*] para ver que N observaciones anteriores influyen sobre la actual, y la autocorrelación parcial (PACF)[*imagen 4*] para comprobar si existe un impacto directo de una observación N .

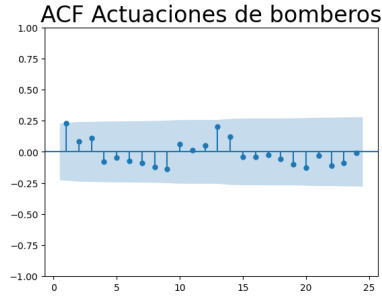


FIGURA 3. imagen
Autocorrelación Simple

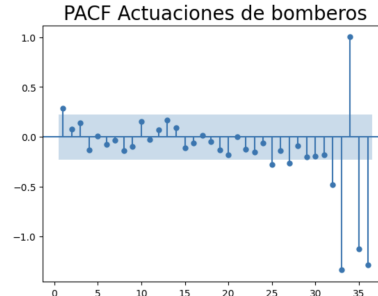


FIGURA 4. imagen
Autocorrelación Parcial

Los resultados obtenidos no sugieren la existencia de una correlación significativa con los meses anteriores, salvo a los 34 meses en la autocorrelación parcial. Este es un comportamiento extraño que, tras ser analizado, se ha visto que los datos de cada 34 meses comparten picos de bajadas y subidas en el número de intervenciones.

3.2. Técnica de suavizado, Holt-Winters. Como primera aproximación se ha probado a realizar predicciones mediante el método de Holt-Winters.

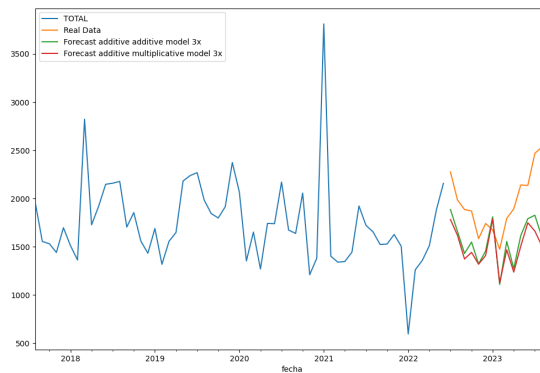


FIGURA 5. Predicción Holt-Winters sobre los bomberos

Como se ha comentado en el notebook, tras probar con distintas distribuciones, las mejores predicciones obtenidas[*imagen 5*] no dan una mala primera aproximación,

pues somos capaces de reconstruir la tendencia que deberían de seguir los datos, aunque con un considerable margen de error.

3.3. Modelos AR y MA. Para poder probar los modelos ARMA, se ha buscado cuales son los mejores valores de $AR(p)$ y $MA(q)$ para nuestra serie temporal. Usaremos la autocorrelación parcial para presuponer los posibles valores de p , pero también se realizarán diversas pruebas dependiendo de los valores de LLF , AIC , BIC y $ERROR$ obtenidos, centrándonos en el error.

En cuanto a la obtención de los valores q en MA nos centraremos en aquellos donde el p-valor sea menor a 0.05, aunque no se descarta que se prueben valores muy cercanos como 0.06. A continuación mostramos las gráficas con las predicciones obtenidas al usar los mejores modelos AR y MA que hemos encontrado (serie en *negro* y *amarillo* respectivamente):

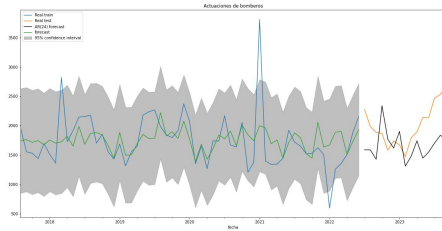


FIGURA 6. imagen
Predicción AR(24)

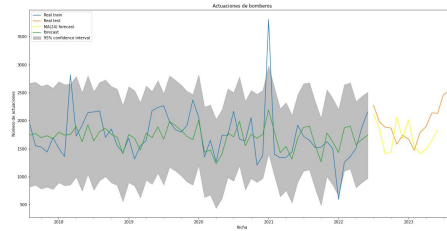


FIGURA 7. imagen
Predicción MA(24)

3.4. Modelos ARMA y ARIMA. Una vez obtenidos los mejores valores de AR y MA, se ha probado a hacer una predicción de los años 2023 y 2024 con el modelo ARMA, véase [imagen 8].

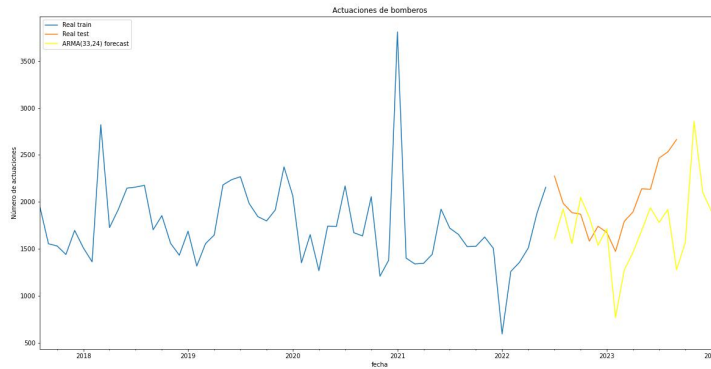


FIGURA 8. Modelo ARMA(33,24) sobre la serie de los Bomberos

Aunque no se hayan realizado integraciones, se ha decidido probar con ARIMA para ver si se obtienen mejores predicciones. Como se puede ver en [imagen 9],

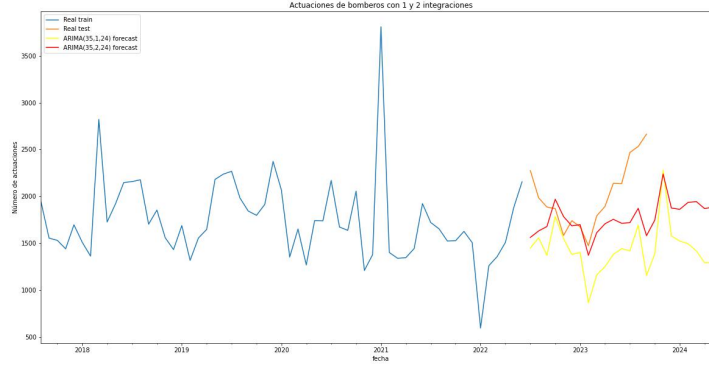


FIGURA 9. Modelo ARIMA(33,n,24) sobre la serie de los Bomberos

La estimación empleando ARIMA con dos integraciones han dado mejores aproximaciones a los datos reales de test, aunque no ha sido capaz de predecir la tendencia ascendente que se ha dado a partir de julio, lo que es esperable pues es un caso atípico.

3.5. Pruebas con redes LSTM. También se decidió realizar diversas pruebas con redes LSTM, pero no se han obtenido buenos resultados. Uno de los motivos por el que creemos que los resultados no son satisfactorios es por la falta de datos, ya que para los bomberos tenemos un total de 74 meses. Que divididos en conjuntos de train y test nos lleva a un total de 59 y 15 observaciones respectivamente.

Debido a dicha falta de datos, aumentar la complejidad de la red hace muy propenso el hecho de producirse overfitting, y tomando medidas para reducir dicho problema como el uso de capas dropout no se han conseguido mejores resultados. Por todo ello, se ha decidido descartar esta aproximación.

4. ANÁLISIS DEL CUERPO DE POLICÍA

La distribución de los datos de la policía [imagen 10] tiene una mayor similaridad con los datos del SAMUR que con los datos de los bomberos, pero aun así son diferentes.

Hay una tendencia descendente durante el año 2017 hasta mediados de 2018, que puede deberse a la puesta en marcha de los planes *Domus y Habitat*[10], que redujeron considerablemente el índice de criminalidad en hogares en múltiples municipios de la Comunidad de Madrid.

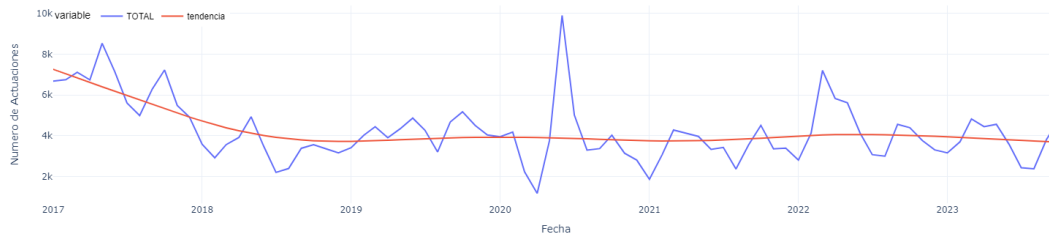


FIGURA 10. Actuaciones de la Policía en el periodo 2017-2023

Hay dos repuntes anómalos que destacan sobre el resto:

- **Junio de 2020**, mes que coincide con la época del confinamiento por el Covid-19. Se puede observar que antes de repunte hay un mínimo histórico entre los meses de marzo y mayo. Cuando se inició el confinamiento el 14 de marzo. Se registra un mínimo global en abril, durante Semana Santa, momento que coincide con la paralelización de actividades no esenciales. En junio se pasó a la fase 0 de la pandemia, donde se permitió volver a realizar actividad física, lo que incrementó el número de denuncias e infracciones hasta alcanzar un máximo histórico[11].
- **Marzo de 2022**: Periodo en el que se incrementó la vigilancia en sitios públicos e inspecciones a locales de ocio, en parte debido a la preocupación de actuaciones de bandas juveniles organizadas.

Excluyendo los casos atípicos, a partir del año 2019 la tendencia se mantiene estable. A diferencia de los bomberos que tienen subidas en verano, el número de incidencias de los policías suele tener 2 repuntes por año, y tiende a caer durante el periodo de agosto para volver a subir en septiembre [12].

4.1. Estudio de Estacionalidad y Correlaciones. La prueba de Dickey Fuller indica un p-valor por debajo de 0.05, por lo que al igual que los bomberos sugiere que la serie es estacionaria.

Tanto la autocorrelación simple [imagen 11] como la parcial [imagen 12] no indican una correlación elevada entre los datos, pero se da cierta importancia a los sucesos del mes previo. También observamos un caso similar al estudio anterior, en el que *PACF* destaca la observación de cada 35 meses.

4.2. Modelos AR y MA. La obtención de los valores p y q siguen la misma dinámica que anteriormente. Para los valores de AR se priorizan aquellos que minimizan el error y para los valores de MA, los que tengan un p-valor por debajo de 0.05. Los resultados obtenidos no son los mejores, pero se aprecia que los modelos son capaces de aproximar la tendencia general de subidas y bajadas en sus respectivos periodos del año.

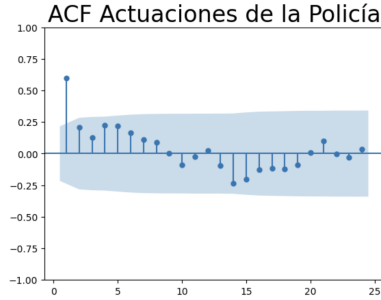


FIGURA 11. imagen
Autocorrelación Simple

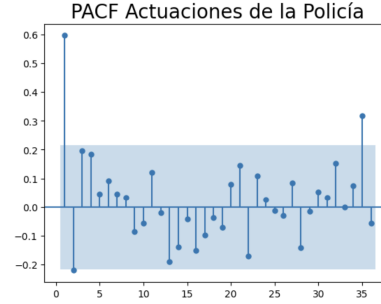


FIGURA 12. imagen
Autocorrelación Parcial

A continuación se muestran dos gráficas con los resultados obtenidos al tratar de usar modelos $AR[imagen\ 13]$ y $MA[imagen\ 14]$ para predecir el número de actuciones del cuerpo de policía (serie *amarilla*):

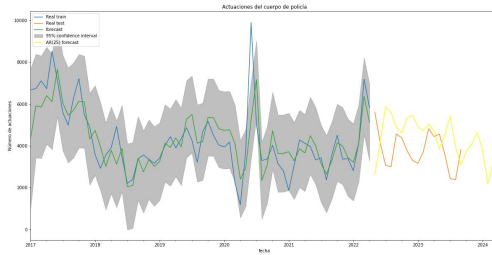


FIGURA 13. imagen
Predicción $AR(25)$

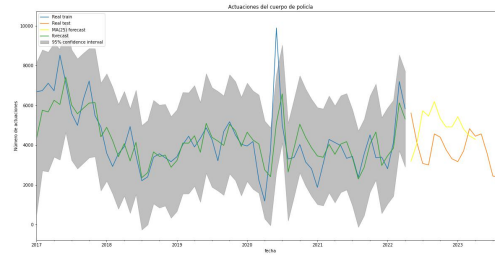


FIGURA 14. imagen
Predicción $MA(25)$

4.3. Pruebas ARMA. Las predicciones realizadas con ARMA no son particularmente destacables, pero el modelo es capaz de aproximar la tendencia general[imagen 15]. Se generan bajadas a mitad de año y subidas a principios y finales. Dicha bajada está ligeramente desplazada, aproximándola ligeramente al principio del verano, cuando suele ser en agosto. Esto, podemos suponer que se debe a la influencia de los meses de covid y el año 2017, que son casos anómalos.

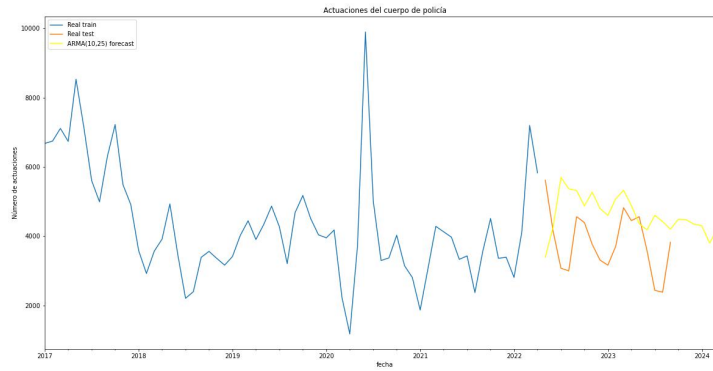


FIGURA 15. Predicciones ARMA(10,25) sobre los datos de la policía

5. ANÁLISIS DEL SAMUR

A diferencia de las series de la policía y los bomberos, el SAMUR tiene una distribución más consolidada.

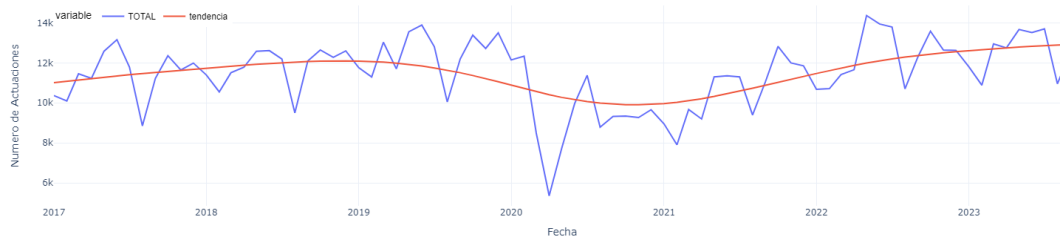


FIGURA 16. Intervenciones del SAMUR en el periodo 2017-2023

Los mínimos anuales en número de intervenciones suelen darse en febrero y agosto, mientras que los máximos en junio y octubre. Por supuesto, debido a la pandemia, el año 2020 es un caso atípico al resto, y cuya distribución es parecida a la del cuerpo de policía, aunque a diferencia de este, no hubo un repunte tan pronunciado tras el mínimo histórico.

5.1. Estudio de Estacionalidad y Correlaciones. En este caso la prueba de Dickey Fuller sí ha dado un p-valor mayor a 0.05, por lo que no se puede rechazar la hipótesis nula de que la serie no es estacionaria. Por tanto, se han realizado dos integraciones hasta obtener un p-valor que nos permita sugerir que la serie es estacionaria.

Al haber calculado los retornos, se ha realizado la autocorrelación parcial y simple tanto de los datos originales[imagen 17][imagen 18] como de los retornos[imagen 19][imagen 20]:

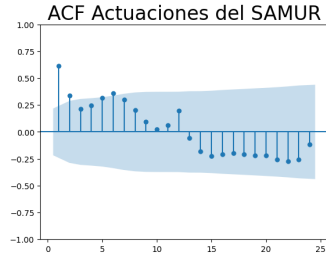


FIGURA 17. imagen Autocorrelación Simple

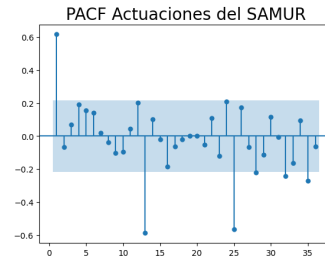


FIGURA 18. imagen Autocorrelación Parcial

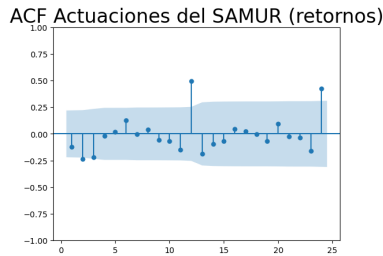


FIGURA 19. imagen Autocorrelación Simple (Con retornos)

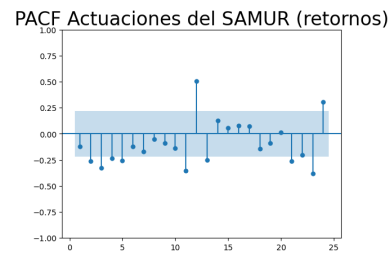


FIGURA 20. imagen Autocorrelación Parcial (Con retornos)

Como se ha podido intuir visualizando la gráfica [imagen 16], se puede observar con las integraciones como los datos de cada 12 meses están correlacionados, ya que el comportamiento del SAMUR es el más regular de los tres cuerpos.

5.2. Modelos AR y MA. Las pruebas con los modelos AR y MA tienden a reducir la varianza de las predicciones sobre el test real. A pesar de ello, son capaces de prever correctamente cuando se van a producir los cambios en aumentos y descensos de incidencias, al menos de forma aproximada.

A continuación hemos obtenido dos gráficas que muestran los resultados alcanzados al tratar de usar modelos AR y MA para predecir el número de actuaciones del SAMUR (serie *amarilla*):



FIGURA 21. figure Predicción AR(17)



FIGURA 22. figure Predicción MA(32)

5.3. Pruebas ARIMA. Se han probado distintos valores de p y q sobre modelos ARIMA, tanto con 1 como con 2 integraciones. Como norma general ARIMA con 1 integración proporciona mejores resultados. Aunque al igual que con AR y MA, se tiende a subestimar la varianza de los datos, por lo que las subidas y bajadas son menos pronunciadas [*imagen 23*].

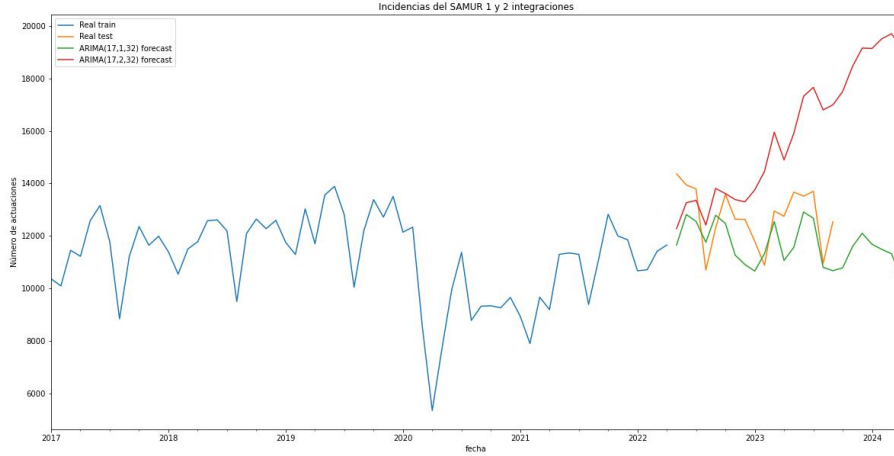


FIGURA 23. Prueba ARIMA sobre los datos del SAMUR

6. INCORPORACIÓN DE VARIABLES EXÓGENAS AL CUERPO DE BOMBEROS

Una vez que hemos obtenido las predicciones de los cuerpos del SAMUR y la Policía, las utilizaremos como variables exógenas para ver qué impacto pueden tener en las intervenciones de los bomberos. Para ello se empleará SARIMAX [*imagen 24*] [13], que nos permite agregar modelos externos a la modelización de la serie y se incorporarán variables seasonal en caso de que se consideren oportunos.

Realizando una comparativa de las diferentes predicciones [*imagen 24*], podemos concluir que el cuerpo de policía no tiene un impacto positivo como variable exógena sobre la predicción del cuerpo de bomberos. Lo que es esperable pues las incidencias de los dos cuerpos no suelen estar correlacionadas. En cambio, el SAMUR sí consigue ajustar mejor las predicciones, aunque la diferencia entre las predicciones usando únicamente el cuerpo de bomberos y las predicciones con el uso del SAMUR como variable exógena no son muy notables.

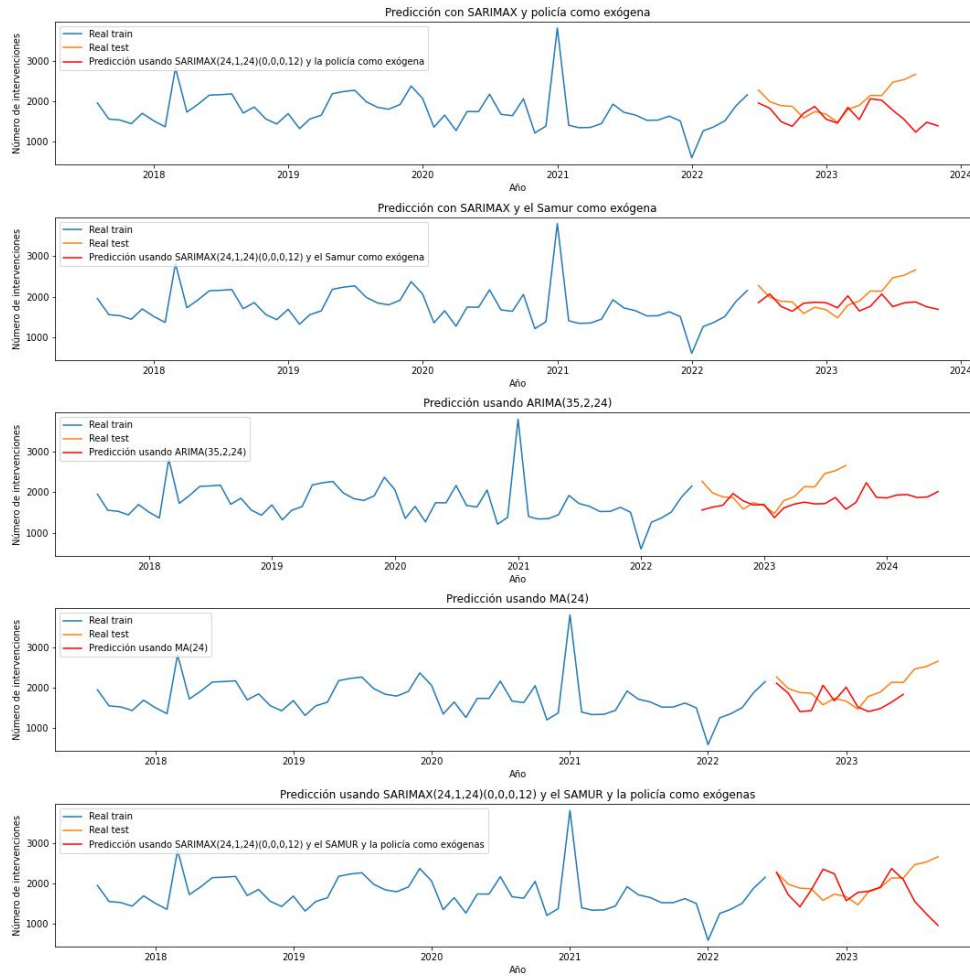


FIGURA 24. Resultados de las predicciones

CUADRO 1. Comparativa de resultados (error)

<i>Modelo</i>	<i>MAE</i>	<i>MSE</i>	<i>RMSE</i>
<i>SARIMAX(policia_exógena)</i>	365,39	282607,48	531,61
<i>SARIMAX(Samur_exógena)</i>	318,86	153313,32	391,55
<i>ARIMA</i>	361,20	220006,98	469,05
<i>MA</i>	398,99	219592,39	468,61
<i>SARIMAX(Samur_y_policia_exógenas)</i>	494,80	426328,59	652,94

La comparativa de errores[[cuadro 1](#)] indican que emplear SARIMAX(24,1,24)(0,0,0,12) y el Samur como variable exógena proporcionan una mayor precisión. Como indica el MAE(*Mean Absolute Error*), se produce un error inferior al resto de modelos con una fluctuación media de 318.86 intervenciones. Además de obtener un

MSE(*Mean Squared Error*) significativamente menor, un estadístico que se enfoca a penalizar las predicciones con errores grandes.

Como la interpretabilidad de MSE puede ser difusa por la propia forma que tiene el estadístico de calcular sus valores, también se ha proporcionado los resultados de RMSE(*Root Mean Squared Error*), pues tratan de escalar los valores de MSE a los datos originales.

También se han realizado pruebas con modelos autoarima/autosarimax para comprobar si hay valores que no se han tenido en cuenta y puedan mejorar las predicciones. Los resultados no han sido mejores que los que se han obtenido previamente [imagen 24], por tanto se han descartado.

En caso de querer considerar una posible predicción válida escogeríamos SARI-MAX(24,1,24)(0,0,0,12) y los datos del SAMUR como exógena, o la modelización usando ARIMA(35,2,24), ya que son los que mejor se aproximan a los resultados que estamos buscando. Consiguiendo asemejarse a la tendencia de los datos de test, por lo menos en cuanto a los aumentos y descensos importantes.

Cabe mencionar que 2023 es un año cuyos meses no están siguiendo la misma tendencia del resto de años (excluyendo casos atípicos como la Filomena). Está presentando una subida continuada en el número de intervenciones hasta los últimos meses disponibles, desde febrero hasta septiembre, cuando lo habitual es que se produzca un repunte de incidencias a partir de abril hasta agosto coincidiendo con la época más calurosa del verano, seguido de un descenso durante el otoño e invierno. Este aumento continuo no se ha debido a un foco específico como en casos anteriores, sino que ha sido por el conjunto total de incidencias que se han generado, tanto adversidades climatológicas como incidentes diversos[14].

REFERENCIAS

- [1] Joaquín Gayoso Cabada. *Documentación de la asignatura de Descubrimiento de Conocimiento en Datos Complejos*.
- [2] Alejandro Rasero Junhao Ge. *Repositorio de Github del proyecto*. URL: <https://github.com/Junhao42/Estudio-y-optimizacion-de-las-actuaciones-de-bomberos-de-la-comunidad-de-Madrid>.
- [3] Ayuntamiento de Madrid. *Actuaciones del Cuerpo de Bomberos*. URL: <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=fa677996afc6f510VgnVCM1000001d4a900aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [4] Ayuntamiento de Madrid. *Policía Municipal. Datos estadísticos actuaciones Policía Municipal*. URL: <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=bffff1d2a9fdb410VgnVCM200000171f5a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [5] Ayuntamiento de Madrid. *Activaciones del SAMUR-Protección Civil*. URL: <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=50d7d35982d6f510VgnVCM1000001d4a900aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [6] David G. Castillejo. *REGIONAL/ 'Filomena'*. URL: <https://noticiasparamunicipios.com/portada/p-regional/regional-filomena-multiplico-por-4-las-intervenciones-de-bomberos-en-enero-de-2021/>.
- [7] Lega21. *Maximo histórico en las actuaciones de bomberos por filomena'*. URL: <https://leganesactualidad.es/el-cuerpo-de-bomberos-de-la-comunidad-de-madrid-realizo-4-962-acciones-en-enero-de-2021-por-filomena/>.
- [8] Cuerpo de bomberos de la comunidad de Madrid. *Informe del cuerpo de bomberos en 2018*. URL: https://www.comunidad.madrid/sites/default/files/memoria_2018_1_0_0.pdf.
- [9] Soporte de Minitab 21. *Interpretar todos los estadísticos y gráficas para Prueba de Dickey-Fuller aumentada*. URL: <https://support.minitab.com/es-mx/minitab/21/help-and-how-to/statistical-modeling/time-series/how-to/augmented-dickey-fuller-test/interpret-the-results/all-statistics-and-graphs/>.
- [10] Ministerio de política territorial y memoria democrática. *Balance de la Criminalidad 2017 de la Comunidad de Madrid*. URL: https://mpt.gob.es/va/delegaciones_gobierno/delegaciones/madrid/actualidad/notas_de_prensa/notas/2018/03/2018-03-3.html.
- [11] Policía Municipal de Madrid. *Balance Covid-19 Fases 0 y 1*. URL: <https://www.madrid.es/UnidadesDescentralizadas/Policia/Publicaciones/RevistaCompromisoyParticipacion/Ficheros/ficheros2020/Junio2020.pdf>.
- [12] Portal web del Ayuntamiento de Madrid. *La Policía Municipal reforzó en 2022 las inspecciones a locales de ocio, la vigilancia en los colegios y el*

- control de la seguridad vial*. URL: <https://www.madrid.es/portales/munimadrid/es/Inicio/Actualidad/Noticias/La-Policia-Municipal-refuerzo-en-2022-las-inspecciones-a-locales-de-ocio-la-vigilancia-en-los-colegios-y-el-control-de-la-seguridad-vial/?vgnextfmt=default&vgnextoid=610ed57de5f56810VgnVCM2000001f4a900aRCRD&vgnnextchannel=a12149fa40ec9410VgnVCM100000171f5a0aRCRD>.
- [13] Joaquín Amat Rodrigo y Javier Escobar Ortiz. *Modelos ARIMA y SARI-MAX con Python*. URL: <https://cienciadedatos.net/documentos/py51-modelos-arima-sarimax-python>.
- [14] Diario de Madrid. *Balance de la actividad, durante 2023, del cuerpo municipal de emergencias*. URL: [https://diario.madrid.es/blog/notas-de-prensa/bomberos-de-madrid-llevo-a-cabo-una-media-de-80-intervenciones-diarias-en-2023-y-rebajo-su-tiempo-de-respuesta-hasta-los-8-minutos-y-31-segundos/#:~:text=Intervenciones%20en%202023&text=Lo%20m%C3%A1s%20habitual%20para%20el,de%20193%20de%20sus%20jornadas\)..](https://diario.madrid.es/blog/notas-de-prensa/bomberos-de-madrid-llevo-a-cabo-una-media-de-80-intervenciones-diarias-en-2023-y-rebajo-su-tiempo-de-respuesta-hasta-los-8-minutos-y-31-segundos/#:~:text=Intervenciones%20en%202023&text=Lo%20m%C3%A1s%20habitual%20para%20el,de%20193%20de%20sus%20jornadas)..)