

# Análisis Completo de la Serie Temporal de los Afiliados a la Seguridad Social

Juan Ignacio Sampere, Jovan Pomar y Gonzalo Jaén

## Resumen

En este análisis completo de las series temporales de los afiliados a la seguridad social, examinamos patrones históricos, tendencias y variaciones en los datos para obtener información sobre las dinámicas sociales y económicas. Utilizamos técnicas estadísticas y computacionales para identificar comportamientos estacionales, predecir tendencias futuras y evaluar indicadores clave que apoyen la toma de decisiones y la asignación de recursos. A través de métodos deterministas aseguramos precisión y robustez en las predicciones. Este trabajo nos permite comprender cambios demográficos, optimizar programas sociales y contribuir a la sostenibilidad financiera del sistema en beneficio de la sociedad.

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Aplicación Empírica</b>	<b>5</b>
2.1. Análisis Exploratorio de los datos . . . . .	5
2.1.1. Serie cogida en su totalidad . . . . .	5
2.1.2. Serie tomada a trozos . . . . .	8
2.2. Metodología . . . . .	10
2.2.1. Preliminares . . . . .	10
2.2.2. Filtro serie total . . . . .	11
2.2.3. Filtro <i>subserie</i> <sub>2</sub> . . . . .	13
2.2.4. Filtros definitivos y comparación de parámetros . . . . .	13
2.3. Resultados . . . . .	15
2.3.1. Interpretaciones . . . . .	15
2.3.2. Predicciones . . . . .	16
<b>3. Conclusiones</b>	<b>18</b>

# 1. Introducción

La crisis financiera de 2008 fue un evento de gran impacto en la economía global, con profundas repercusiones en mercados, empleo y sistemas de seguridad social. Su origen puede rastrearse a Estados Unidos y se debió a una combinación de factores como la liberalización financiera y la proliferación de productos financieros complejos, cuya regulación resultó insuficiente. El detonante de la crisis fue la burbuja inmobiliaria estadounidense, que creció en gran parte debido a prácticas de préstamo de alto riesgo, conocidas como hipotecas *subprime*, otorgadas a personas con baja capacidad de pago. Estas hipotecas se combinaron y vendieron en forma de valores respaldados por hipotecas, distribuyendo el riesgo entre inversores de todo el mundo. Cuando los precios de las viviendas comenzaron a caer, millones de personas incumplieron sus pagos, generando pérdidas masivas para bancos y empresas de inversión como *Lehman Brothers*. Esto produjo una cadena de quiebras bancarias y, al restringir el crédito global, se desencadenó una crisis económica en numerosos países.

La crisis de 2008 tuvo efectos drásticos en el empleo y el sistema de seguridad social. Al reducirse la actividad económica y restringirse el acceso al crédito, muchas empresas despidieron a parte de sus plantillas para reducir costes. Este aumento en el desempleo afectó el número de personas afiliadas a la seguridad social, ya que muchos dejaron de cotizar al perder su empleo formal. En España, los niveles de desempleo alcanzaron cifras récord, y la afiliación a la seguridad social disminuyó considerablemente durante los años siguientes a la crisis. Esta pérdida de cotizantes, además de reducir los ingresos para la seguridad social, planteó desafíos para la sostenibilidad del sistema de protección social, que depende de las aportaciones de trabajadores activos. El análisis de la evolución del número de afiliados a la seguridad social permite observar cómo una crisis económica puede influir a largo plazo en el empleo y en la estabilidad del sistema de seguridad social.

Para estudiar estos efectos, hemos recopilado una **serie temporal de los afiliados a la seguridad social en España** que abarca desde la década de 1980 hasta 2019. Estos datos, obtenidos de fuentes oficiales como el Instituto Nacional de la Seguridad Social\*, cubren el periodo anterior a la crisis de 2008, la etapa de recesión, y los años de recuperación económica. A continuación, mostramos un gráfico de la serie temporal completa, que permite observar la tendencia general de la afiliación a lo largo de los años, con los efectos visibles de la crisis de 2008 y la recuperación posterior (Figura 1).

Este análisis se organiza en dos partes. En primer lugar, clasificamos la serie temporal (sólo tomaremos los valores de 2009 a 2018), descomponiéndose en componentes de tendencia, estacionalidad y ruido, lo cual nos ayuda a identificar patrones específicos y entender los efectos de las distintas etapas económicas. En segundo lugar, aplicamos varios modelos predictivos (y decidiremos cuál es el que mejor se adecua a la serie) para prever la evolución futura de los afiliados a la seguridad social, lo cual es relevante para orientar políticas y decisiones en la gestión del sistema

---

\*Los datos provienen de la web del Ministerio de Empleo y Seguridad Social de España (<http://www.empleo.gob.es/index.htm>) en el apartado Estadísticas / Mercado de Trabajo / Movimiento Laboral Registrado / Principales Series.

de protección social en el contexto actual. La motivación detrás de la elección de analizar la serie del número de afiliados a la seguridad social en España es porque actualmente el sistema de seguridad social enfrenta serios desafíos de sostenibilidad. Con el envejecimiento de la población, la disminución de la natalidad y las fluctuaciones en el mercado laboral, los ingresos del sistema están bajo presión mientras aumenta la demanda de prestaciones. Nuestro estudio pretende aportar una perspectiva histórica y predictiva sobre cómo ha evolucionado el número de afiliados, para entender mejor los efectos a largo plazo de crisis económicas como la de 2008 y cómo estos patrones pueden ayudar a orientar políticas efectivas en el futuro.

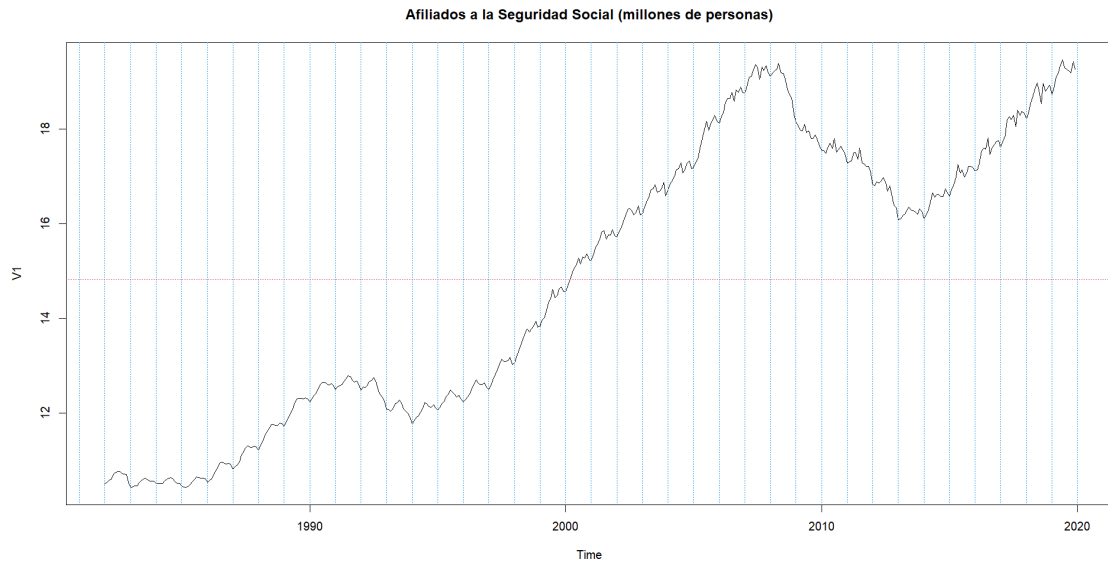


Figura 1: Serie temporal del número de afiliados a la Seguridad Social en España desde 1980 a 2019

## 2. Aplicación Empírica

En este apartado, clasificaremos nuestra serie, ajustaremos varios filtros con los que daremos predicciones futuras y daremos nuestras interpretaciones de cuáles podrían ser las causas del comportamiento de la serie.

### 2.1. Análisis Exploratorio de los datos

#### 2.1.1. Serie cogida en su totalidad

La serie contiene un total de 120 observaciones, iniciando en enero de 2009 y finalizando en diciembre de 2018. Durante este periodo, se observa un comportamiento que refleja tanto las fluctuaciones económicas como las recuperaciones y caídas en el empleo. El resumen estadístico de la serie es el siguiente: la media del número de afiliados se sitúa en 17.36 millones de personas, con una desviación estándar de 0.76 millones, lo que indica una variabilidad moderada en el número de afiliados a lo largo del tiempo.

El valor mínimo registrado es de 16.08 millones, coincidiendo con los puntos más bajos de la crisis económica que se reflejan en la gráfica entre los años 2011 y 2013. Por otro lado, el valor máximo es de 18.97 millones, alcanzado en los últimos meses de 2018, lo cual sugiere una recuperación importante del mercado laboral hacia el final del periodo de análisis.

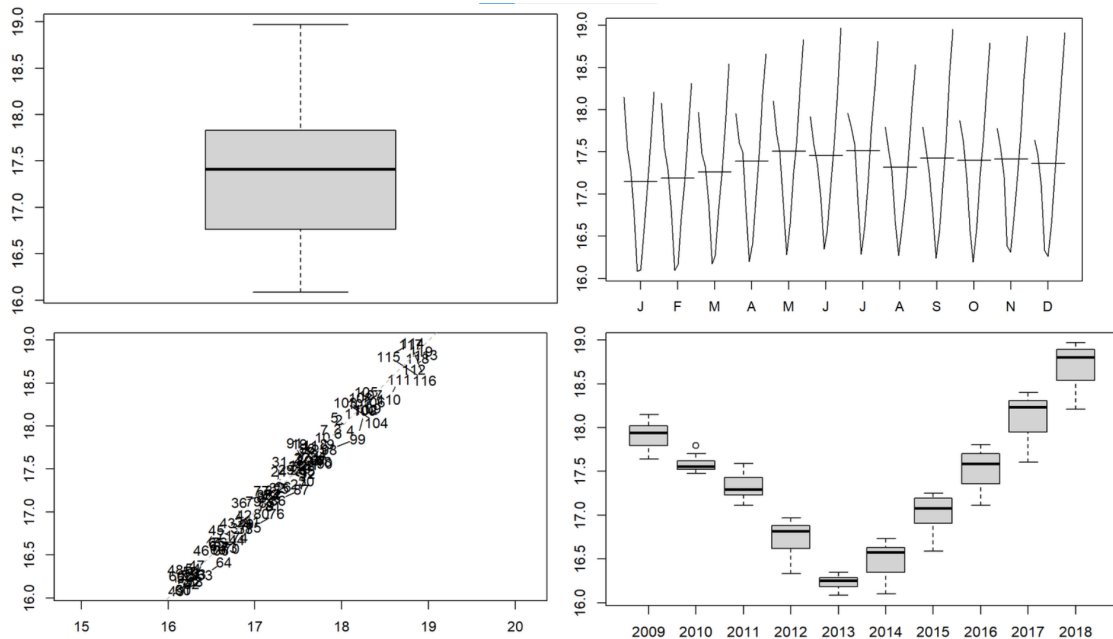


Figura 2: *Boxplot*, *Boxplot* mensual, *Lagplot* y *Boxplot* anual de la serie

El análisis gráfico de la serie temporal se enriquece al observar el *monthplot*, el *lagplot* y el *boxplot* anual (Figura 2). El *monthplot* revela un comportamiento estacional, con picos notables en abril y julio y descensos en agosto y diciembre, indicando variaciones típicas del mercado laboral relacionadas con la temporalidad. El *lagplot* muestra una clara autocorrelación positiva, evidenciada por la alineación diagonal

de los puntos, lo que sugiere que los valores mensuales están fuertemente relacionados con sus predecesores. Por último, el *boxplot* anual ilustra cómo la varianza de la serie cambió con el tiempo: entre 2011 y 2013, durante la crisis, los valores fueron más bajos y menos dispersos. A partir de 2014, se observa un aumento progresivo en la mediana y una mayor dispersión, reflejando una recuperación sostenida del empleo hasta 2018. Este comportamiento sugiere un mercado laboral que se estabiliza y crece, con variaciones más marcadas en los años de recuperación.

El filtro lineal centrado de rango  $2k^\dagger$  compensado en los extremos es una técnica útil en R para suavizar una serie temporal y aproximar su componente de tendencia:

$$Y_t = \frac{1}{2} \frac{1}{2k} X_{t-k} + \sum_{i=-(k-1)}^{k-1} \frac{1}{2k} X_{t-i} + \frac{1}{2} \frac{1}{2k} X_{t+k}$$

Cuando se aplica este filtro y se grafica la serie (Figura 3), se obtiene un gráfico que muestra la serie original con una línea roja (la tendencia) que destaca el movimiento general sin el ruido de la estacionalidad y otras fluctuaciones de corto plazo.

Esta línea de tendencia permite ver claramente cómo la serie ha evolucionado a lo largo de los años, evidenciando periodos de descenso, como durante la crisis económica, y de crecimiento sostenido en la recuperación posterior. La estacionalidad, que se refiere a los patrones que se repiten periódicamente dentro de un año, se hace más evidente cuando se elimina la tendencia, lo que permite enfocarse en las variaciones cíclicas mensuales. Este análisis es clave para identificar cómo los afiliados a la Seguridad Social tienden a comportarse de manera similar en los mismos meses de diferentes años, un aspecto importante para la planificación y el análisis predictivo.

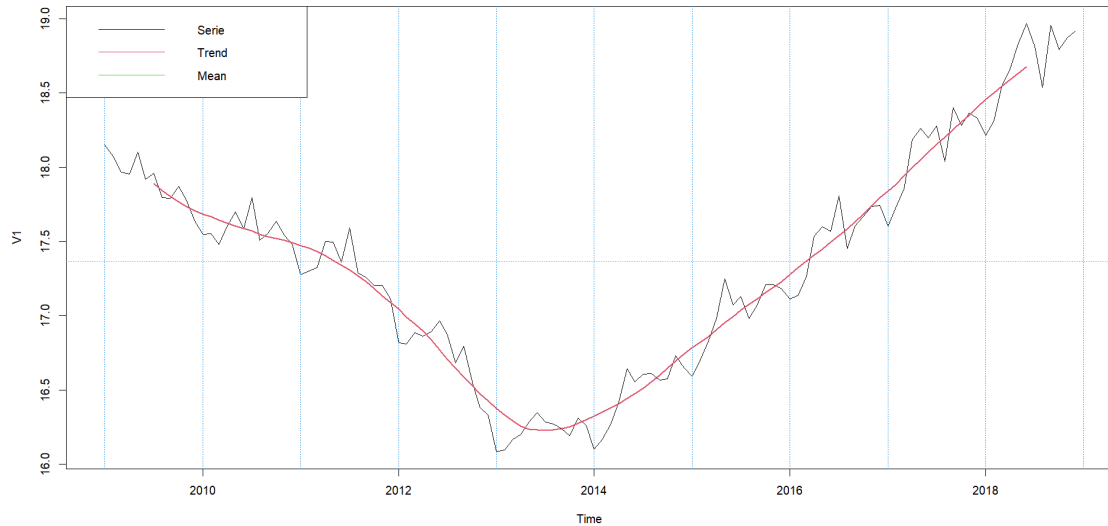


Figura 3: Gráfico de la serie con su tendencia aproximada por el filtro lineal centrado  $2k$  compensado en los extremos

Cuando calculamos los índices estacionales (Figura 4) de la diferencia entre la serie y su tendencia aproximada anteriormente, se destacan las fluctuaciones esta-

---

<sup>†</sup>Cogemos la  $k$  de tal forma que  $2k = \text{frecuencia}$ , en nuestro caso la frecuencia es de 12 por ser datos mensuales así que  $k = 6$ .

cionales alrededor de un valor medio. Este gráfico sugiere una estacionalidad que parece mantenerse relativamente constante a lo largo de los años, lo que apunta a una posible estacionalidad aditiva.

También podríamos ver que los meses de abril a julio suelen estar por encima del promedio de la serie sin tendencia, mientras que meses como de enero y febrero tienden a estar por debajo. Este comportamiento refuerza la idea de que la estacionalidad tiene un patrón cíclico definido.

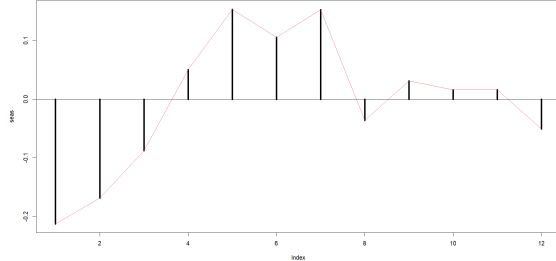


Figura 4: Índices estacionales

Al descomponer la serie (Figura 5), observamos claramente la división en sus componentes: tendencia, estacionalidad y ruido. En nuestra serie, la componente estacional parece mantenerse estable en amplitud, lo que sugiere que la serie podría clasificarse como AAA (Aditiva, Aditiva, Aditiva). Sin embargo, si se notara una variación en la magnitud proporcional al nivel de la serie, podríamos considerar un modelo AAM (Aditiva, Aditiva, Multiplicativa). Esta descomposición visual nos ayuda a identificar la estructura subyacente de la serie sin sacar conclusiones definitivas aún.

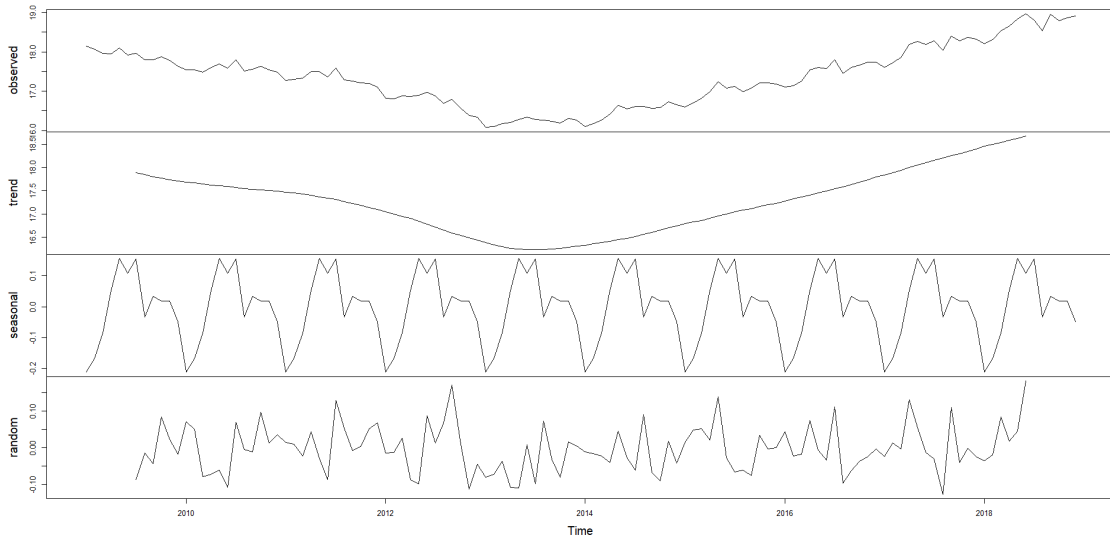


Figura 5: Descomposición aditiva de la serie

Finalmente, dividiremos la serie en dos periodos (por ejemplo, 2009-2013 y 2014-2018) que permitirán estudiar las diferencias en estacionalidad y tendencia en cada subperiodo. En el periodo de crisis, las fluctuaciones estacionales pueden ser más sutiles debido a la baja en la afiliación total, mientras que en los años de recuperación, estas variaciones pueden intensificarse junto con el crecimiento de la tendencia. Esta comparación nos da más contexto para entender cómo la estacionalidad y la varianza de la serie se comportan en distintos escenarios económicos.

### 2.1.2. Serie tomada a trozos

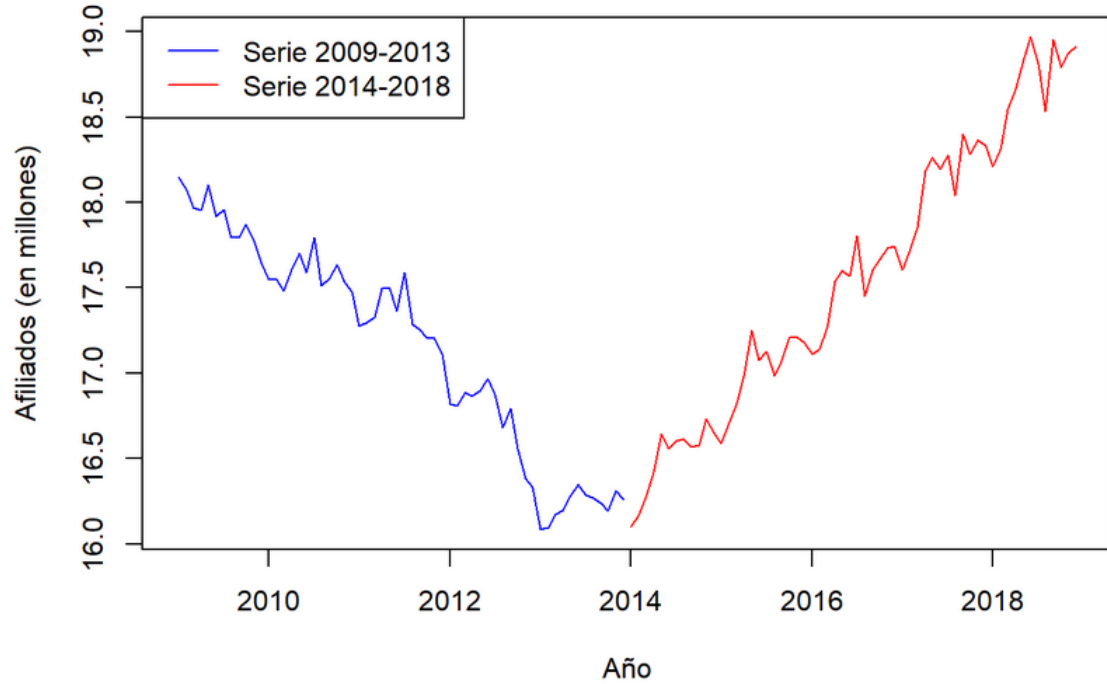


Figura 6: Número de afiliados a la Seguridad Social con sus subseries visibles

Al dividir la serie de afiliados a la Seguridad Social en dos subseries ( $subserie_1$  de 2009-2013 y  $subserie_2$  de 2014-2018, Figura 6), podemos observar cómo los patrones de tendencia y estacionalidad varían entre un periodo de crisis económica y uno de recuperación. Para ambas subseries, seguimos el análisis realizado previamente con la serie total para entender mejor sus características y el posible impacto en las predicciones.

Estadísticos	$subserie_1$	$subserie_2$
Min.	16.08	16.10
1st Qu.	16.65	16.94
Median	17.29	17.58
Mean	17.16	17.57
3rd Qu.	17.61	18.28
Max.	18.15	18.97

Cuadro 1: Resumen de las dos subseries.

En la  $subserie_1$ , nuestro análisis (Figura 7) revela una varianza relativamente baja, lo cual es consistente con un periodo de contracción económica donde los valores de afiliación se mantienen en niveles bajos y estables. Destaca estacionalidad menos pronunciada; los picos y valles mensuales, aunque presentes, son más uniformes en amplitud. La tendencia muestra un descenso sostenido, y al las fulctuaciones siguen



siendo visibles pero de menor magnitud. Los índices estacionales, confirman que los patrones cíclicos persisten, aunque parecen más suaves durante esta etapa.

En la *subserie*<sub>2</sub>, los resultados son notablemente diferentes (Figura 7). La varianza tiene rango más amplio, evidenciando un periodo de recuperación económica donde la afiliación total crece de forma sostenida. Los patrones estacionales son más marcados, con picos notables en meses como abril y julio, y caídas en agosto y diciembre, lo que coincide con la recuperación del mercado laboral y su dinamismo. Además la tendencia es claramente linealmente ascendente con unas fluctuaciones estacionales más visibles, sugiriendo una estacionalidad más intensa que en los años de crisis. Los índices estacionales en este periodo muestran que las variaciones mensuales son más significativas, especialmente en los meses de mayor actividad económica. La descomposición de la serie evidencia una componente de tendencia fuerte y estable, con una estacionalidad aditiva.

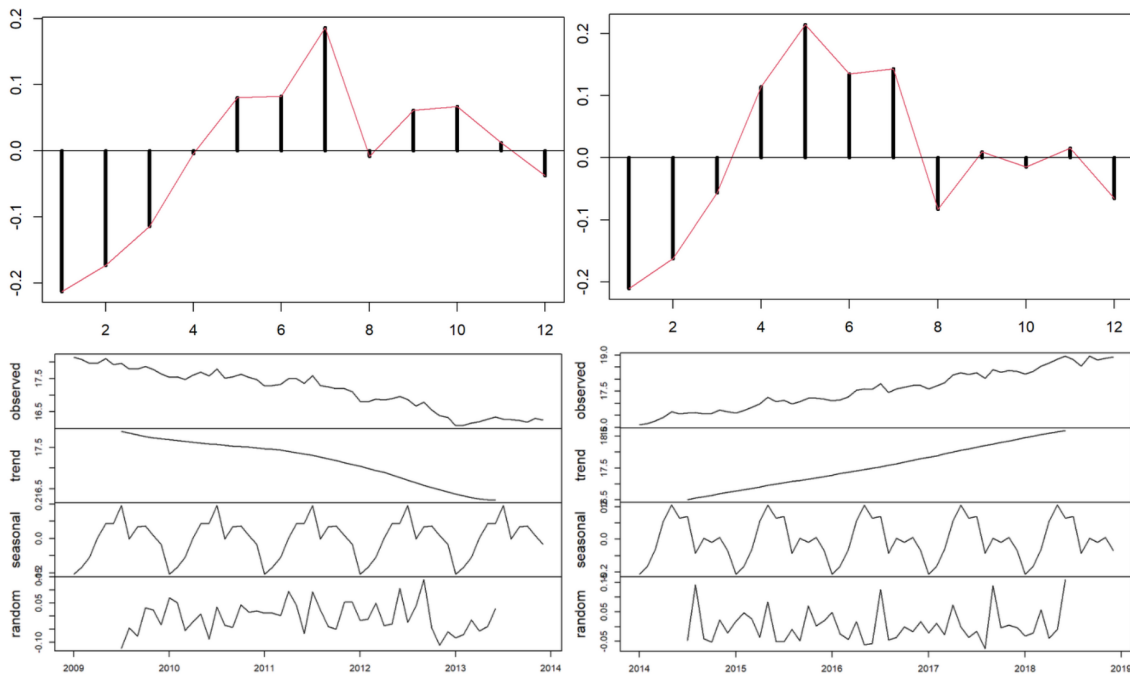


Figura 7: Índices estacionales y descomposición de la *subserie*<sub>1</sub> (izquierda) y de la *subserie*<sub>2</sub> (derecha)

En ambas subseries, podemos decir que la estacionalidad está presente. En la *subserie*<sub>1</sub> dada su heterocedasticidad se clasificaría como una serie de tipo 6 (AAM<sup>‡</sup>); mientras que con la clara homocedasticidad de los datos desde 2014-2018 se deduce que la *subserie*<sub>2</sub> es de tipo 4 (AAA<sup>‡</sup>).

La amplitud de las fluctuaciones es más prominente en el periodo de 2014-2018, lo que puede influir en la elección de modelos predictivos. Esta comparación sugiere que las predicciones realizadas en cada subperiodo podrían variar en precisión debido a las diferencias en la estabilidad y amplitud de la estacionalidad. La división permite observar cómo la serie ha evolucionado de manera más específica.

<sup>‡</sup>Notación ETS (*Error Trend and Seasonality*)

## 2.2. Metodología

### 2.2.1. Preliminares

Una vez clasificadas las series en nuestro análisis exploratorio toca ver cual de los filtros que hemos aprendido se ajusta mejor a las series. Como dichos filtros nos ayudan después a dar predicciones escogeremos un mejor filtro para la serie total y otro filtro para la *subserie*<sub>2</sub>.

En ambos casos procederemos de la misma manera. Hemos escogido tres diferentes funciones de errores que optimizaremos mediante los parámetros de tres filtros distintos. Las funciones de los errores son las siguientes:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad \text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

En cuánto a los filtros a comparar, al estar trabajando con series con tendencia y estacionalidad, usaremos el filtro de Holt Winters para series con estacionalidad aditiva (tipo 4, AAA) y el mismo filtro adaptado a series con estacionalidad multiplicativa (tipo 6, AAM). También queremos probar el filtro de Holt para las series con tendencia pero sin estacionalidad (tipo 2, AAN) asegurándonos que necesitamos asumir ésta última componente.

Antes de pasar a dar los resultados de los ajustes aplicados recordemos ciertos puntos sobre los filtros Holt y Holt Winters.

- **Tipo 2:** con tendencia y sin estacionalidad (AAN, Holt)

Ecuación de predicción:  $\hat{X}_{T+h|T} = a_T + hb_T$

Ecuación para el suavizado del nivel:

$$a_t = \alpha X_t + (1 - \alpha)(a_{t-1} + b_{t-1})$$

Ecuación para el suavizado de la pendiente (tendencia):

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)a_{t-1}$$

Valores iniciales:  $\alpha, a_0, \beta, b_0$     Parámetros a estimar:  $\alpha, \beta$ .

- **Tipo 4:** con tendencia y con estacionalidad aditiva de orden  $m$ <sup>§</sup>(AAA, Holt Winters)

Ecuación de predicción:  $\hat{X}_{T+h|T} = a_T + hb_T + S_{T+h-m}$

Ecuación para el suavizado del nivel:

$$a_t = \alpha(X_t - S_{t-m}) + (1 - \alpha)(a_{t-1} + b_{t-1})$$

Ecuación para el suavizado de la tendencia:

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)a_{t-1}$$

Ecuación para el suavizado de la estacionalidad:

$$S_t = \gamma(X_t - a_{t-1} - b_{t-1}) + (1 - \gamma)S_{t-m}$$

Valores iniciales:  $\alpha, a_0, \beta, b_0, \gamma, S_{-m+1}, \dots, S_0$     Parámetros a estimar:

$\alpha, \beta, \gamma$ .

---

<sup>§</sup>Nuestro patrón se repite cada año, por tanto como los datos son mensuales, en nuestro caso  $m = 12$ .

- **Tipo 6:** con tendencia y con estacionalidad multiplicativa de orden  $m$  (AAM, Holt Winters)

Ecuación de predicción:  $\hat{X}_{T+h|T} = (a_T + hb_T)S_{T+h-m}$

Ecuación para el suavizado del nivel:

$$a_t = \alpha \frac{X_t}{S_{t-m}} + (1 - \alpha)(a_{t-1} + b_{t-1})$$

Ecuación para el suavizado de la tendencia:

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$$

Ecuación para el suavizado de la estacionalidad:

$$S_t = \gamma \frac{X_t}{a_{t-1} + b_{t-1}} + (1 - \gamma)S_{t-m}$$

Valores iniciales:  $\alpha, a_0, \beta, b_0, \gamma, S_{-m+1}, \dots, S_0$  Parámetros a estimar:

$\alpha, \beta, \gamma$ .

Los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  en el suavizamiento exponencial Holt Winters determinan el peso que se asigna a los datos recientes frente a los históricos al estimar el nivel, la tendencia y la estacionalidad de una serie temporal.

Un  $\alpha$  alto (cercano a 1) indica que el modelo responde rápidamente a cambios en el nivel de la serie, lo que es útil para series volátiles o con cambios abruptos. En contraste, un  $\alpha$  bajo da más peso a los datos históricos, produciendo un ajuste más estable y menos reactivo, ideal para series suaves con pocos cambios significativos.

El parámetro  $\beta$  controla el ajuste de la tendencia. Un  $\beta$  alto permite que el modelo se adapte rápidamente a cambios recientes en la pendiente de la serie, siendo útil cuando hay aceleraciones o desaceleraciones frecuentes. Por otro lado, un  $\beta$  bajo implica que el modelo prefiere mantener una tendencia más estable y se ajusta lentamente, algo adecuado para series con patrones de tendencia persistentes y consistentes. Cuando se incluye estacionalidad, el parámetro  $\gamma$  regula cómo se ajusta el modelo a los patrones estacionales. Un  $\gamma$  alto responde rápidamente a cambios en la estacionalidad, mientras que uno bajo suaviza esos patrones, adaptándose lentamente a las variaciones.

La interacción entre estos parámetros define cómo de reactivo o estable es el modelo en su respuesta a los cambios en la serie. Valores altos para todos los parámetros hacen que el modelo sea más sensible a cambios en nivel, tendencia y estacionalidad, pero pueden amplificar el ruido, haciéndolo menos estable. Valores bajos, en cambio, generan un modelo conservador que da más peso a los patrones históricos, siendo más adecuado para series con comportamientos predecibles.

### 2.2.2. Filtro serie total

Nuestra hipótesis sobre la presencia de estacionalidad en la serie ha sido confirmada tras el análisis de los modelos aplicados. Al comparar los resultados obtenidos con los diferentes enfoques de Holt Winters, es evidente que debemos optar por un modelo que incluya un componente estacional, ya sea aditivo o multiplicativo (sus errores son significativamente menores, Figura 8). Sin embargo, dado que la serie presenta una varianza no constante a lo largo del tiempo, el modelo más adecuado será aquel que pueda manejar esta característica de manera efectiva.

---

¶El color rojo indicaría las predicciones de los dos siguientes años para cada caso que comentaremos en el apartado de Resultados.

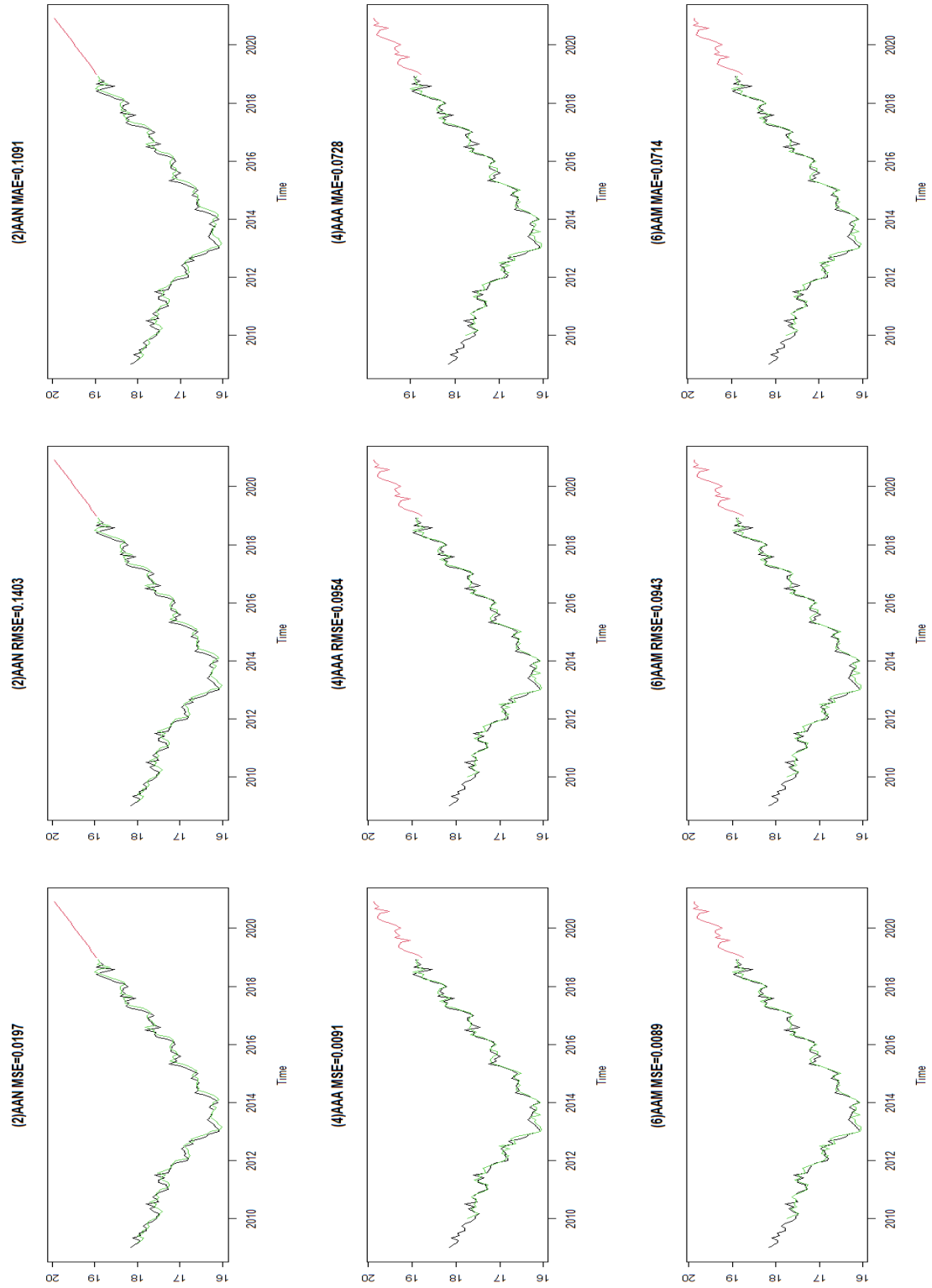


Figura 8: Filtros ajustados a la serie, en verde, con sus valores correspondientes optimizando la tres distintas funciones de error.¶

Entre los modelos evaluados, tanto el de estacionalidad aditiva como el de estacionalidad multiplicativa requieren estimar el mismo número de parámetros, pero el enfoque multiplicativo tiene la ventaja de adaptarse mejor a series en las que la magnitud de los patrones estacionales varía proporcionalmente al nivel. Por esta razón, hemos decidido quedarnos con el filtro de Holt Winters que utiliza una estacionalidad multiplicativa, ya que ofrece un ajuste más preciso a la estructura de la serie y es más robusto frente a los cambios en la varianza observada.

### 2.2.3. Filtro $subserie_2$

Con las observaciones anteriores ya hemos asumido que hay estacionalidad. Por ello sólo aplicaremos los dos filtros de Holt Winters de tipo 4 y de tipo 6. Observando los valores de los errores, ambos son muy parecidos (Figura 9). Nos podríamos sentir tentados a favorecer el filtro de tipo 6 por tener errores ligeramente más pequeños, tal y como hemos hecho en el caso anterior de la serie total. Sin embargo aquí no lo haremos porque tenemos muy clara la clasificación de la  $subserie_2$ , que, por su varianza constante posee estacionalidad aditiva. Así pues nos quedamos con el filtro Holt Winters AAA de tipo 4.

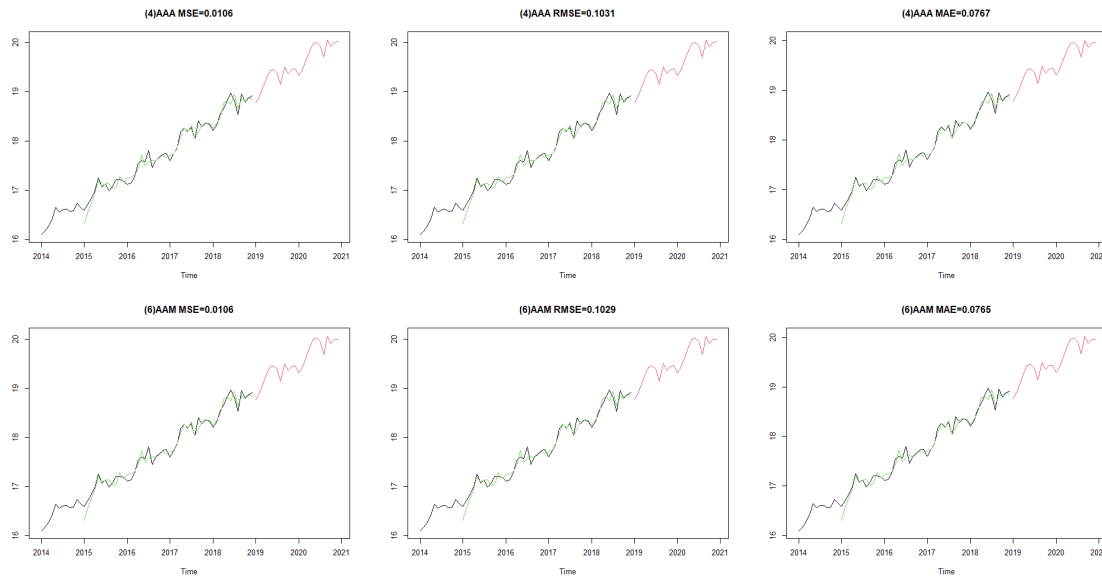


Figura 9: Filtros ajustados a la  $subserie_2$ , en verde, con sus valores correspondientes optimizando la tres distintas funciones de error.

### 2.2.4. Filtros definitivos y comparación de parámetros

Resumen, hemos escogido para la serie total aplicar el filtro Holt Winters de estacionalidad multiplicativa y para la  $subserie_2$  el filtro de Holt Winters de estacionalidad aditiva. Véase en la siguiente tabla los parámetros de los anteriores filtros, una vez automatizado el proceso de ajustes con las funciones en R ya implementadas:

Caso Aditivo <i>subserie<sub>2</sub></i>		Caso Multiplicativo <i>serie</i>	
Parámetros	Valor	Parámetros	Valor
$\alpha$	0.398768	$\alpha$	0.459852
$\beta$	0.006909952	$\beta$	0.2211541
$\gamma$	0.8318089	$\gamma$	0.6695764
Coefficientes	Valor	Coefficientes	Valor
$a$	18.939477834	$a$	18.90088998
$b$	0.045844671	$b$	0.04027791
$s_1$	-0.205941480	$s_1$	0.99030386
$s_2$	-0.131913706	$s_2$	0.99316396
$s_3$	0.010918522	$s_3$	0.99945814
$s_4$	0.147269229	$s_4$	1.00682888
$s_5$	0.254258534	$s_5$	1.01153584
$s_6$	0.236028780	$s_6$	1.00887741
$s_7$	0.119722119	$s_7$	1.00500737
$s_8$	-0.15995390	$s_8$	0.99137794
$s_9$	0.151661909	$s_9$	1.00689302
$s_{10}$	-0.033837776	$s_{10}$	0.99921164
$s_{11}$	0.004403121	$s_{11}$	1.00201231
$s_{12}$	-0.031193464	$s_{12}$	1.00021340

En el caso aditivo  $\alpha = 0,398768$  indica un valor intermedio que sugiere una actualización moderada del nivel basado en los datos recientes. La baja magnitud de  $\beta = 0,006909$  indica que la tendencia se ajusta lentamente, lo que implica que la serie tiene una evolución de tendencia suave. Este valor alto de  $\gamma = 0,8318089$  sugiere que los ajustes estacionales tienen un gran peso, lo cual es esperable al tener la serie componentes estacionales marcados.

En el caso multiplicativo  $\alpha = 0,459852$  es un poco mayor, lo que implica una mayor sensibilidad al nivel actual de la serie.  $\beta = 0,2211541$  también mayor que en el caso aditivo, lo que indica que el componente de tendencia cambia más rápido. Y,  $\gamma = 0,6695764$  aunque es alto, tiene menos peso relativo comparado con el caso aditivo, lo que implica que la estacionalidad influye de manera diferente.

En cuanto a los coeficientes estacionales, en el caso aditivo, los valores más extremos son los coeficientes  $s_1 = -0,205941480$ ,  $s_2 = -0,131913706$ ,  $s_8 = -0,159095390$  (negativamente) y  $s_5 = 0,254258534$ ,  $s_6 = 0,236028780$  (positivamente). Estos sugieren que hay meses con un comportamiento significativamente diferente (pico o caída notable). En el caso multiplicativo, la estacionalidad parece más regular, con valores más cercanos a 1, aunque el índice de los coeficientes más significativos coinciden con los del caso aditivo, que corresponden a los meses de enero, febrero, mayo, junio y agosto.

## 2.3. Resultados

### 2.3.1. Interpretaciones

Una vez analizada nuestra serie temporal mediante una metodología basada en las matemáticas estadísticas nos complacería dar las interpretaciones correspondientes bajo un punto de vista realístico. Para ello recopilamos las conclusiones más destacables hasta el momento.

El análisis de la serie temporal que muestra el número de afiliados a la Seguridad Social desde 2009 a 2018 se puede separar de dos trozos, tal y como hemos hecho anteriormente, una serie de 2009 a 2013 (*subserie<sub>1</sub>*) y otra serie de 2014-2018 (*subserie<sub>2</sub>*).

La *subserie<sub>1</sub>* tiene una tendencia decreciente, se observa una clara disminución del número de afiliados que seguramente es una consecuencia de la crisis del 2008. Durante esta etapa muchos trabajos fueron destruidos lo que provocó un aumento del paro en nuestro país. También esta serie muestra una varianza no constante, las fluctuaciones en el número de afiliados son amplias y no siguen un patrón predecible. Esta alta volatilidad es un indicador de la incertidumbre económica, los cambios drásticos en el mercado laboral, la inestabilidad de la política laboral y los efectos de la crisis económica global.

En cambio la *subserie<sub>2</sub>* muestra una tendencia creciente lineal, el número de afiliados crece constantemente a lo largo de los próximos cinco años pasando de 16.2 millones a 18.8 millones de personas. Esto refleja una recuperación económica estable. Y, es posible que los esfuerzos de esta recuperación, las políticas de empleo o la estabilidad política hayan contribuido a este crecimiento. Además estamos ante un claro caso de homocedasticidad lo que sugiere que el mercado laboral ha logrado una cierta estabilidad después de la crisis. Esto también puede reflejar la normalización de las condiciones laborales y una recuperación gradual en los niveles de empleo.

A lo largo de ambas series tenemos fluctuaciones estacionales regulares, un patrón similar que se repite a lo largo de todos los años, ya que los flujos de empleo tienden a tener ciclos anuales, influenciados por la contratación estacional, las vacaciones o los ciclos económicos. En la *subserie<sub>2</sub>* la afiliación muestra picos y valles en los meses que tradicionalmente tienen mayor o menor contratación, sin los choques repentinos que caracterizaron la *subserie<sub>1</sub>*. Estudiemos estos índices estacionales para sacar conclusiones más precisas.

Vemos que de diciembre a marzo el número de afiliados sociales tiende a bajar cada año mientras que de abril a julio crece y en agosto vuelve a haber un índice negativo. Esto se puede deber a varios factores, el más importante siendo que España es un país con alta dependencia del turismo. El turismo da lugar a una mayor dependencia del empleo estacional, sobre todo en verano. El decrecimiento de diciembre a marzo es debido al fin de contratos temporales, la baja de actividad en sectores estacionales y una menor actividad económica general. El crecimiento de abril a julio es debido al incremento de empleo en sectores turísticos, agrícolas y de servicios, junto con contratación estacional para el verano. Agosto sería por tanto el fin de los contratos de verano y la constancia de septiembre a noviembre representaría la vuelta a trabajos estables o en caso de los jóvenes adultos la vuelta a los estudios.

### 2.3.2. Predicciones

Ahora bien, todos estos puntos nos han ayudado a entender nuestra serie y poder ajustar el mejor filtro posible, explicado en el apartado anterior de metodología. Hemos decidido, debido a la clasificación de las series y a los errores, escoger el filtro de las series de tipo AAM para la serie y el filtro de las series de tipo AAA para la *subserie*<sub>2</sub>. Con estos dos filtros obtendremos dos predicciones distintas para los siguientes dos años<sup>‡</sup>, 2019 y 2020 que pretendemos comparar entre ellos y con los valores reales obtenidos por la Seguridad Social.

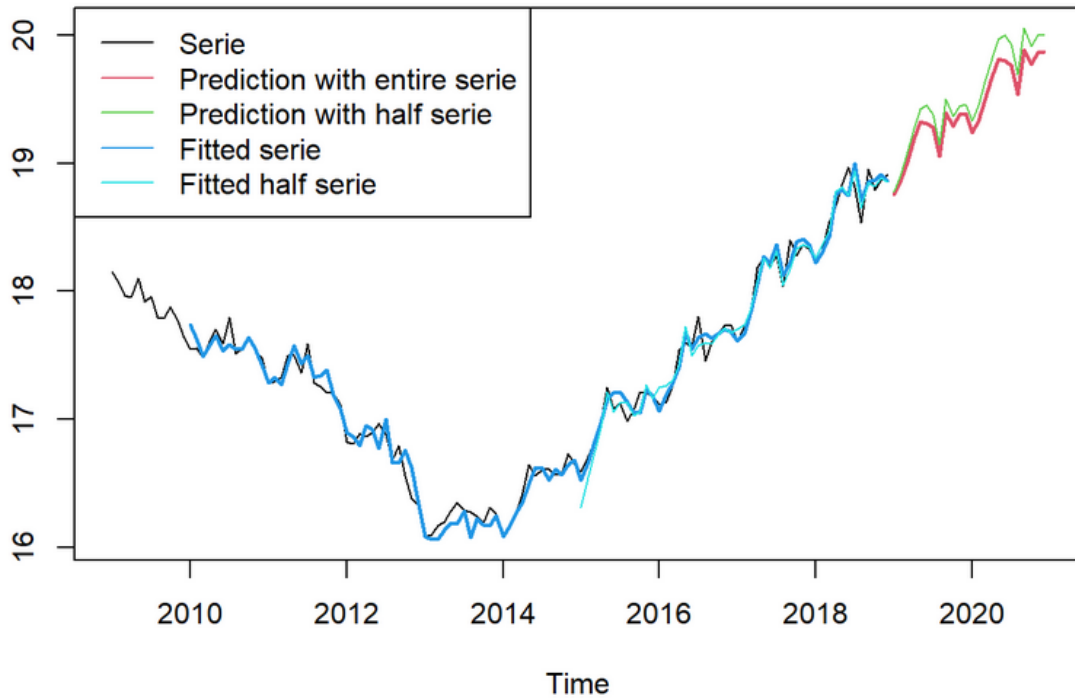


Figura 10: Gráfico de la serie con los filtros definitivos Holt Winters y sus predicciones

Vemos que ambas predicciones continúan con el mismo patrón repetitivo estacional (Figura 10). La predicción obtenida con la totalidad de la serie da valores menores que la predicción hecha con la segunda parte de la serie. Podríamos ir incluso más lejos y atrevernos a decir que la primera predicción es una translación vertical negativa de la segunda predicción. Que la segunda predicción sea más elevada que la primera tiene sentido puesto que, al sólo considerar la parte creciente de la tendencia de la serie, no arrastra los valores de la primera parte, aquella dónde la tendencia es decreciente.

Sin embargo, calculando la suma de los errores cuadráticos de cada serie con respecto a los valores reales del año 2019 nos da  $err_{AAM} = 0,01058221$  y  $err_{AAA} = 0,01610507$ . Esto nos hace pensar que considerar la serie en su totalidad nos hubiese ayudado a predecir mejor los valores futuros. Es decir que en nuestro caso siguen siendo relevantes los valores de la *subserie*<sub>1</sub>.

<sup>‡</sup>Aquí el  $h_{step} = 24$  puesto que en 2 años hay 24 meses.



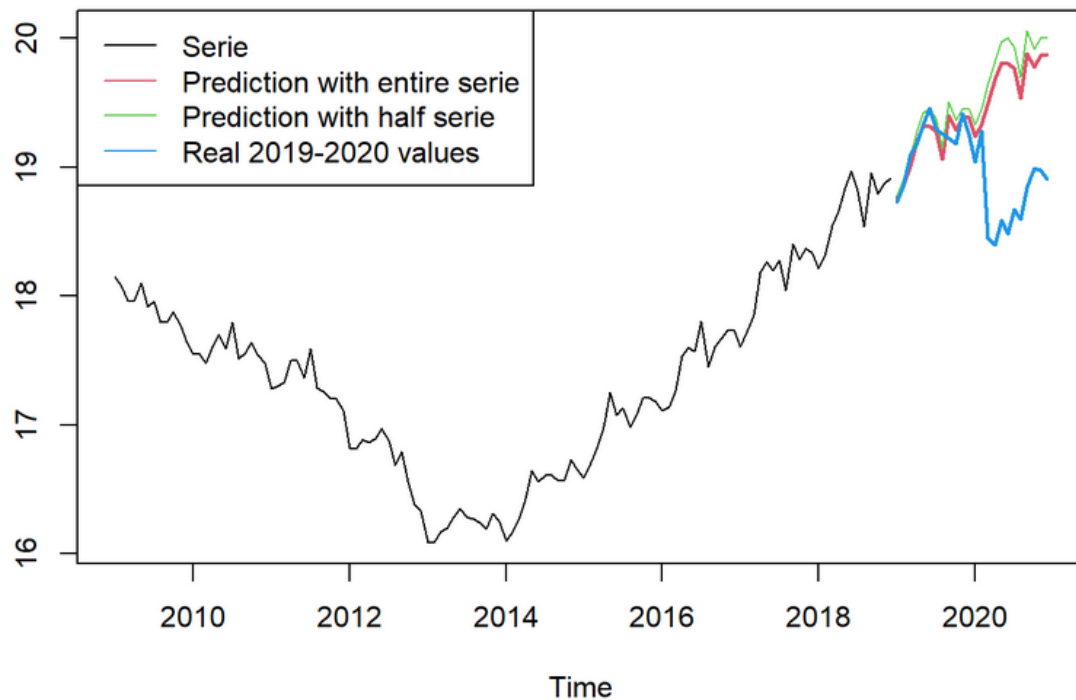


Figura 11: Gráfico de comparación entre las predicciones de los filtros escogidos y los verdaderos valores futuros

Ahora bien, observamos que si comparamos con los valores reales del siguiente año (Figura 11), nuestras predicciones no hubiesen servido y dan mucho error. No han predicho una súbita bajada en la primera mitad del año 2020 del número de afiliados a la Seguridad Social. La causa de este súbito cambio de patrón es la aparición del fatídico coronavirus. La pandemia que provocó este virus afectó directamente a los afiliados haciendo que bajase de 900.000 en un sólo mes.

Los filtros que hemos usado, basan sus predicciones en la extrapolación de patrones históricos: estacionalidad, tendencia y nivel. Es decir, las predicciones se construyen bajo la premisa de que el futuro será una extensión de los patrones observados en el pasado. Por ello mismo, al no considerar eventos externos presentan limitaciones significativas para predecir eventos extraordinarios como crisis económicas o pandemias.

### 3. Conclusiones

El análisis de la serie temporal de afiliados a la Seguridad Social en España (2009-2018) evidencia una tendencia general ascendente, marcada por una caída significativa durante la crisis económica (2009-2013) y una recuperación sostenida a partir de 2014. Durante la crisis, el número de afiliados alcanzó su mínimo histórico de 16.08 millones, con valores bajos y poca dispersión, mientras que en la recuperación se registró un crecimiento continuo hasta alcanzar un máximo de 18.97 millones en 2018, acompañado de una mayor variabilidad. Además, la serie presenta una estacionalidad aditiva consistente, con picos de afiliación en abril y julio, y descensos en agosto y diciembre, asociados a la temporalidad laboral.

La descomposición de la serie muestra tres componentes clave: tendencia, estacionalidad y ruido, destacando una recuperación económica que intensifica las fluctuaciones estacionales. Gráficos como el *boxplot* y el *lagplot* confirman una mayor dispersión y autocorrelación positiva durante la recuperación. Este análisis subraya cómo las dinámicas laborales están influenciadas por factores económicos y estacionales, proporcionando información valiosa para la planificación estratégica y el diseño de modelos predictivos que consideren la variabilidad y estacionalidad de los distintos periodos analizados.

Tras realizar un análisis detallado de nuestra serie temporal utilizando herramientas de estadística matemática, es importante interpretar los resultados desde un punto de vista realista. La serie de afiliados a la Seguridad Social entre 2009 y 2018 se divide en dos etapas clave: la primera, de 2009 a 2013 (*subserie<sub>1</sub>*), marcada por una tendencia decreciente debido a la crisis económica de 2008, con una alta volatilidad que refleja incertidumbre y cambios abruptos en el empleo. Por el contrario, la segunda etapa, de 2014 a 2018 (*subserie<sub>2</sub>*), muestra una tendencia creciente lineal, acompañada de estabilidad en la varianza, lo que indica una recuperación económica y una normalización del mercado laboral. Ambas etapas presentan fluctuaciones estacionales regulares, con patrones predecibles como descensos de diciembre a marzo y aumentos de abril a julio, influenciados por la dependencia del empleo estacional en sectores como el turismo y la agricultura.

Los análisis predictivos basados en estas series sugieren que los patrones estacionales y las tendencias históricas son útiles para predecir valores futuros bajo condiciones normales. Sin embargo, al comparar las predicciones de 2019 y 2020 con los valores reales, se observa que eventos extraordinarios como la pandemia del coronavirus introducen cambios abruptos en la serie que no pueden ser anticipados con los modelos utilizados. Aunque las predicciones basadas en la serie completa (filtro AAM) ofrecen menor error en 2019, ninguna logró prever el impacto drástico de la pandemia en 2020, lo que subraya las limitaciones de los métodos estadísticos basados exclusivamente en patrones históricos. Este análisis enfatiza la necesidad de complementar las metodologías tradicionales con técnicas adaptativas o basadas en escenarios para incorporar posibles eventos externos inesperados.

## Referencias

- [1] George E. P Box. *Time Series Analysis: Forecasting and Control*. Hoboken N.J.: 4th ed. Wiley, 2008.
- [2] Ministerio de Empleo y Seguridad Social. *Datos del número de afiliados a la Seguridad Social*. Accedido: 21 de noviembre de 2024. 1980-2018. URL: <http://www.empleo.gob.es/index.htm>.