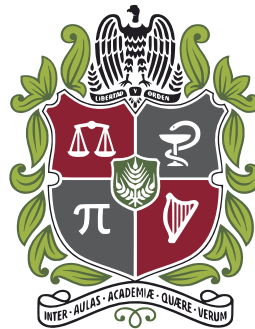


Proyección de poblaciones carcelarias en Colombia

SERGIO DAVID SOLANO BEJARANO
INGENIERO INDUSTRIAL



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
ABRIL DE 2017

Proyección de poblaciones carcelarias en Colombia

SERGIO DAVID SOLANO BEJARANO
INGENIERO INDUSTRIAL

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE
MASTER EN CIENCIAS - ESTADÍSTICA

DIRECTOR
B. PIEDAD URDINOLA CONTRERAS, PH.D.
DOCTOR EN DEMOGRAFÍA

LÍNEA DE INVESTIGACIÓN
DEMOGRAFÍA



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
ABRIL DE 2017

Título en español

Proyección de poblaciones carcelarias en Colombia

Title in English

Prison populations projections for Colombia

Resumen: Se realizan proyecciones de la población carcelaria en Colombia, usando la información disponible para los años 1991-2017. La información publicada periódicamente no incluye las tasas de transición (ingreso y salida) del sistema, por esta razón se eligen tres métodos que permiten realizar la proyección a partir de la población observada. Los métodos utilizados son: modelos demográficos para poblaciones pequeñas, modelos ARIMA, y modelos Estado-Espacio. Se comprara el ajusted de cada modelo, sus ventajas y desventajas.

Abstract: Projections of the prison population in Colombia are made, considering available data from 1991-2017. Monthly released data does not include admission or release rates, so we use three methods that work over the total population. The considered methods are: Demographical methods for subnational populations, ARIMA Models, and State-Space Models. We compare the fit of the three models, their advantages and disadvantages.

Palabras clave: Poblaciones carcelarias, series de tiempo, procesos SARIMA, Modelos Estado Espacio, Poblaciones pequeñas

Keywords: Prison populations, time series, SARIMA processes, State Space Models, Subnational Populations

Nota de aceptación

Trabajo de tesis

Aprobado

“Mención Meritoria o Laureada”

Jurado

Jurado uno

Jurado

Jurado dos

Director

B. Piedad Urdinola

Bogotá, D.C., Abril 19 de 2017

Dedicado a

Este documento está dedicado a Hercilia Bejarano, mi madre.

Agradecimientos

Agradezco a Piedad, paciente directora de tesis, por su colaboración; a mis compañeros de estudio: Julian, Jennifer y Ángela; y a Alanis, por supuesto.

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
1. Antecedentes teóricos	1
1.0.1. Proyecciones de población	1
1.0.2. Proyección de poblaciones pequeñas	2
1.0.3. Aplicaciones nacionales e internacionales	2
1.0.4. Series de tiempo	3
1.0.5. Modelos ARIMA	4
1.0.6. Modelos Estado-Espacio	4
2. La población carcelaria en Colombia 1991 - 2017	6
2.1. Análisis exploratorio	6
2.2. El sistema penitenciario en Colombia	7
2.2.1. Identificación del modelo	7
3. Modelos SARIMA	9
3.1. Identificación del modelo	9
3.2. Estimación de parámetros	13
3.3. Proyección	28
3.4. Conclusiones	30
4. Métodos demográficos para poblaciones pequeñas	32
4.1. Censal ratio	34

4.2. Población privada de la libertad por rango de edad Febrero 2016 a Febrero 2018	36
4.3. Proyecciones Censal Ratio	42
5. Modelos estado-espacio	43
5.1. Selección de Software	43
5.2. Identificación del modelo	43
5.3. Simulación Monte-Carlo	43
5.4. Estimación de parámetros	43
5.5. Proyecciones 2017 - 2020	43
5.6. Conclusiones	43
Conclusiones	44
Trabajo futuro	45
Bibliografía	46

Índice de tablas

3.1. Comportamiento de la ACF y PACF en modelos ARMA(p,q) [19]	12
3.2. Comparación de criterios de información	14
3.3. Parámetros del modelo (3,2,1,0,0,2)	14
3.4. Parámetros del modelo (1,2,1,0,0,2)	14
3.5. Comparación de criterios de información	19
3.6. Parámetros del modelo (0,1,1,0,0,1)	19
3.7. Comparación de criterios de información	24
3.8. Parámetros del modelo (1,1,1,0,0,1)	24

Índice de figuras

2.1. Población privada de la libertad 1991 - 2017	6
2.2. Tasa de encarcelamiento según genero 1991 - 2017	7
2.3. Población carcelaria por situación jurídica	8
3.1. Variación inter-mensual de población carcelaria	9
3.2. Descomposición de la variación inter-mensual de población carcelaria total .	10
3.3. Descomposición de la variación inter-mensual de población carcelaria total .	11
3.4. Descomposición de la variación inter-mensual de población carcelaria total .	11
3.5. Autocorrelación parcial de la variación inter-mensual	12
3.6. Diagnóstico del sarima (1,1,1)(0,0,2), población total	15
3.7. Diagnóstico del sarima (1,1,1)(0,0,1), población total	16
3.8. Diagnóstico del sarima (3,2,1)(0,0,2), población total	17
3.9. Diagnóstico del sarima (1,2,1)(0,0,2), población total	18
3.10. Diagnóstico del sarima (1,1,1)(0,0,1), población sindicada	20
3.11. Diagnóstico del sarima (2,1,1)(0,0,1), población sindicada	21
3.12. Diagnóstico del sarima (0,1,2)(0,0,1), población sindicada	22
3.13. Diagnóstico del sarimfa (0,1,2)(0,0,0), población sindicada	23
3.14. Diagnóstico del sarima (1,1,1)(0,0,1), población condenada	25
3.15. Diagnóstico del sarima (2,1,1)(0,0,1), población condenada	26
3.16. Diagnóstico del sarima (0,1,2)(0,0,1), población condenada	27
3.17. Diagnóstico del sarima (0,1,2)(0,0,0), población condenada	28
3.18. Proyección de la población carcelaria total	29
3.19. Proyección de la población carcelaria sindicada	29
3.20. Proyección de la población carcelaria condenada	30
4.1. Proyección de la población nacional por rango de edad	33

4.2. Piramide poblacional, población privada de la libertad, Junio - 2017	36
4.3. Población privada de la libertad. Febrero-2016 a Febrero-2018	37
4.4. Población privada de la libertad. Febrero-2016 a Febrero-2018	38
4.5. Población privada de la libertad. Febrero-2016 a Febrero-2018	39
4.6. Población privada de la libertad. Febrero-2016 a Febrero-2018	40
4.7. Tasa de específica de encarcelamiento. Junio-2017	41
4.8. Tasa específica de encarcelamiento. Junio-2016, Junio -2017	41
4.9. Población privada de la libertad. 2005-2020	42

CAPÍTULO 1

Antecedentes teóricos

En 2016 Colombia ocupó el puesto catorce entre doscientos cincuenta y un países por el tamaño de su población carcelaria (120 914 hbts.) y el cincuenta y uno según la tasa de encarcelamiento (240 por cada 100.000 hbts). Tasa que pasó de 51,5 en el año 2000 a 240 por cada 100.000 hbts en 2016. Con una ocupación del 154 % de las plazas disponibles, resulta relevante contar con proyecciones de la población carcelaria en el corto, mediano y largo plazo.[11]

1.0.1. Proyecciones de población

“Una estimación poblacional consiste en determinar el tamaño o las características de una población, para el momento actual o para uno anterior, en ausencia de información. Cuando se realizan un conjunto de supuestos sobre el comportamiento de los vitales hacia el futuro, hablamos de proyección, y cuando se escoge un escenario como el más probable, hablamos de pronóstico” [21].

Para incluir la incertidumbre en las proyecciones de población Lee enumera los siguientes métodos [13]:

- El enfoque de escenarios alto, medio y bajo: Asume comportamientos fijos para la fertilidad, la mortalidad y las migraciones durante el período de proyección, basado en algunos supuestos [13].
- Análisis estocásticos
 - Análisis ex-post: consiste en evaluar el error de pronóstico en proyecciones anteriores y aplicarlo a las nuevas proyecciones [13].
 - Simulación estocástica: Permite hacer proyecciones de población, al asignar una distribución de probabilidad a las tasas vitales (mortalidad, natalidad, migraciones) [13].
 - Modelos estocásticos de la tasa de crecimiento: Consiste en estimar la tasa de crecimiento del total de la población; aunque permite estimar intervalos de confianza, no permite separar la proyección de las tasas vitales, ni de las franjas etarias [13].

- Matrices de Leslie con modelos estimados para las tasas vitales: Al usar matrices de Leslie se estima la población por rangos etarios para un instante i , y se calcula la población en el instante $i + 1$ aplicando la natalidad y la mortalidad proyectadas para el período i . Puesto que las series de población carcelaria no se publican separadas por edad, no se abordará esta técnica.

1.0.2. Proyección de poblaciones pequeñas

La proyección de áreas pequeñas es entendida como la proyección a un nivel geográfico menor al nacional. Estas proyecciones pueden incluir, departamentos, ciudades o poblaciones especiales [21]. “Una población especial es un grupo poblacional que se encuentra restringido a un área por una medida administrativa o legislativa. Dentro de los grupos usualmente considerados se encuentran las prisiones, universidades, hospitales e instituciones militares”. Este tipo de población puede tener una estructura etaria y de sexo, y unos vitales diferentes al resto de la población; además no suelen envejecer en el mismo lugar, lo que permite mantener una estructura etaria que no varía a través del tiempo. [21].

1.0.3. Aplicaciones nacionales e internacionales

Las proyecciones de poblaciones carcelarias oficiales analizadas corresponden, en buena parte, a los métodos expuestos en los capítulos anteriores: Proyecciones por escenarios, proyección de la tasa de crecimiento, modelos ARIMA para las tasas de ingreso y salida.

En Colombia (CONPES 3828) se proyectó la población carcelaria usando la tasa media de crecimiento anual (1993-2014) [7]. Esta proyección no tiene en cuenta la incertidumbre asociada con las variaciones aleatorias en las tasas, ni las asociadas a cambios estructurales, ni que son una población especial, con patrones muy distintos a los del resto de la población [21] .

El Reino Unido hasta 2015 realizaba una proyección por escenarios (alto, medio y bajo), año en el cual cambió a un modelo de proyección de la media y su incertidumbre. La incertidumbre se incluyó a través de un análisis ex-post, de la desviación de la proyección en años anteriores [12].

El departamento de Justicia de los Estados Unidos realizó estimaciones de la población carcelaria por estado para el período 2013-2014. Las estimaciones parten del censo de prisiones 1993-2014. Estas proyecciones se puede enmarcar dentro de las proyecciones de áreas pequeñas [15].

El bureau de estadísticas e investigación del crimen en Australia proyecta las tasas de arresto y sentencia usando modelos ARIMA; a partir de estas tasas proyecta la población carcelaria. Estas proyecciones incluyen un período de validación de tres años. Los resultados mostraban que la serie real se encuentra dentro de los intervalos de confianza de la proyección, cercano a la proyección de la media [23].

Blummstein desarrolla un método de proyección basado en los componentes demográficos, tasas específicas de arresto por delito y reincidencias, a partir de estos datos proyecta el tamaño y la composición de las poblaciones [1].

Con los datos libres disponibles en Colombia no se podría utilizar el enfoque ARIMA ni el método de Blummstein, pues las tasas de encarcelamiento y sentencia no se publican. La

tesis busca proponer un método de proyección, para situaciones donde no se cuenta con el registro de los vitales o su equivalente en la población analizada, en este escenario resulta conveniente usar las series ARIMA sobre la serie de población sindicada y población sindicada, y los modelos estado-espacio para la serie bivariada. Se proponen los métodos demográficos de estimación para áreas pequeñas Censal-ratio y Ratio-Correlation, y su contraste con las metodologías ARIMA y Estado-Espacio que a diferencia de los métodos demográficos mencionados incorporan una medida de incertidumbre.

1.0.4. Series de tiempo

Una serie de tiempo es un conjunto de observaciones x_t asociadas a un instante de tiempo t . Es usual referirse como series de tiempo, tanto a las realizaciones x_t como a las variables aleatorias X_t que las generan. [4]

En una regresión lineal clásica, una variable Y es explicada o predicha en función de una variable X . La diferencia entre el valor observado y_i y el valor predicho se suponen provenientes de un proceso aleatorio con media cero. [5]

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (1.1)$$

Donde $\{\epsilon_1, \epsilon_2, \dots\}$ son independientes e idénticamente distribuidas (i.i.d.). Es posible representar una serie de tiempo en esta forma, tomando como variable explicativa el tiempo. Sin embargo, el análisis de series de tiempo se ha desarrollado como un área particular de la estadística, pues es este supuesto no suele cumplirse, enfocándose particularmente en variables aleatorias dependientes, que pueden ser representadas en la forma [4]:

$$Y_i = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + \dots + \beta_{i-1} Y_{i-1} + \epsilon_i \quad (1.2)$$

Para medir la dependencia entre las observaciones recurrimos a la función de autocovarianza y a la función de autocorrelación. La función de autocovarianza se define como:

$$\gamma(h, h+k) = \text{cov}(x_h, x_{h+k}) = E[(x_h - \mu_h)(x_{h+k} - \mu_{h+k})] \quad (1.3)$$

para todo h y k mayor a cero. La autocovarianza mide la dependencia lineal entre dos puntos en la misma serie en diferentes instantes. [19].

La función de autocorrelación se define como [19]:

$$\rho(h, h+k) = \frac{\gamma(h, h+k)}{\sqrt{\gamma(h, h)\gamma(h+k, h+k)}} \quad (1.4)$$

Para poder realizar predicciones sobre el estado futuro de una serie de tiempo, es necesario suponer que el comportamiento de la serie es estable a través del tiempo, incluso con un componente aleatorio. La estacionaridad es el concepto que permite articular esta necesidad. Estamos ante un **proceso estacionario** cuando las variables $\{X_1, \dots, X_k\}$ tienen la misma distribución conjunta que $\{X_{h+1}, \dots, X_{h+k}\}$, para todos los enteros h y k [4].

La estacionaridad débil se presenta cuando $E[X_j]$ y $E[X_j X_{j+h}]$ son independientes de j , es decir: i) Presenta media μ constante e independiente de t ii) La función de autocovarianza

$\gamma(h, h+k)$ depende solamente de la cantidad de pasos que separa las observaciones (k). [19]

1.0.5. Modelos ARIMA

Box & Jenkins proponen una aproximación iterativa de cuatro etapas para la selección de un modelo [2]:

1. Selección de una clase de modelos, con base en la teoría y la práctica.
2. Identificación del modelo, donde se seleccionan un conjunto de parámetros que permitan explicar el sistema con parsimonia.
3. Estimación de parámetros.
4. Chequeo diagnóstico, para detectar fallas en el ajuste. Si se detectan fallas en el ajuste, se regresa al segundo paso.

Los procesos ARMA son procesos aleatorios de la forma [19] :

$$Y_t = \gamma Y_{t-1} + \gamma_2 Y_{t-2} \dots + \gamma_i Y_{t-i} + \epsilon + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \dots \theta_j \epsilon_{t-j} \quad (1.5)$$

A este modelo se le conoce como ARMA(i,0,j).

Los procesos ARIMA resultan al considerar una serie de la forma [19]:

$$Y_t = \alpha + Y_{t-1} \quad (1.6)$$

Tal que el proceso $Y_t - Y_{t-1}$ es un proceso ARMA.

Los criterios para identificar el orden de un ARMA son [3]:

- Cuando la función de autocorrelación (ACF) se reduce progresivamente, y la función de autocorrelación parcial (PACF) no tiene picos en lags luego de p, es un proceso autoregresivo de orden p AR(p)
- Cuando la función de autocorrelación tiene un pico en el lag q, y la función de autocorrelación parcial se reduce progresivamente, es un proceso de media mmovil de orden q MA(q)
- Si ambas funciones se reducen gradualmente, se trata de un ARMA (p,q)

1.0.6. Modelos Estado-Espacio

Tomado de [14] En un modelo estado espacio una serie de tiempo (multiple) observada y_1, \dots, y_t depende de un estado z_t , posiblemente no observado, que se comporta siguiendo un proceso estocástico. La relación entre y_t y z_t está dada por la *ecuación de medida*: [14]

$$y_t = H_t z_t + v_t \quad (1.7)$$

donde H_t es una matriz que puede o no depender del tiempo t y v_t es el error de observación, que se asume usualmente como un proceso de ruido. El vector de estado es generado como:

$$z_t = B_{t-1}z_{t-1} + w_{t-1} \quad (1.8)$$

La matriz B_t es una matriz de coeficientes que puede depender de t y w_t es un proceso de ruido. [14]

Es posible considerar los modelos estado-espacio una generalización de los modelos ARIMA y VARMA

La población carcelaria en Colombia 1991 - 2017

2.1. Análisis exploratorio

El INPEC publica mensualmente la serie población carcelaria. Esta serie contiene la población carcelaria desde 1991, separada por situación jurídica (condenados, sindicados) y género.

La población carcelaria total entre 1991 y 2017 se ha cuadruplicado, al pasar de 32.036 a 128.125 internos. Ver figura 2.1. Aunque la mayoría de los internos son hombres, la población carcelaria femenina ha crecido a un ritmo aún más acelerado, al quintuplicar su población entre 1991 y 2017 (pasa de 1633 personas a 7800).

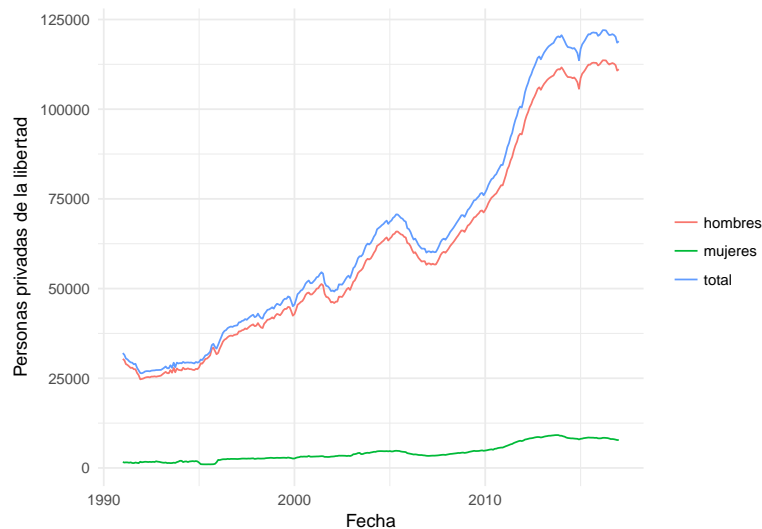


FIGURA 2.1. Población privada de la libertad 1991 - 2017

Fuente: INPEC Elaboración propia

El incremento en la población carcelaria podría tomarse como un efecto del crecimiento de la población colombiana. Para validar este supuesto calculamos la tasa de encarcelamiento, que mide la cantidad de personas encarceladas por cada cien mil habitantes. Este

indicador pasó de 92 personas por cada cien mil habitantes en enero de 1991 a 242 en enero de 2016. Tal incremento se puede ver tanto en hombres como en mujeres. Ver figura 2.2.

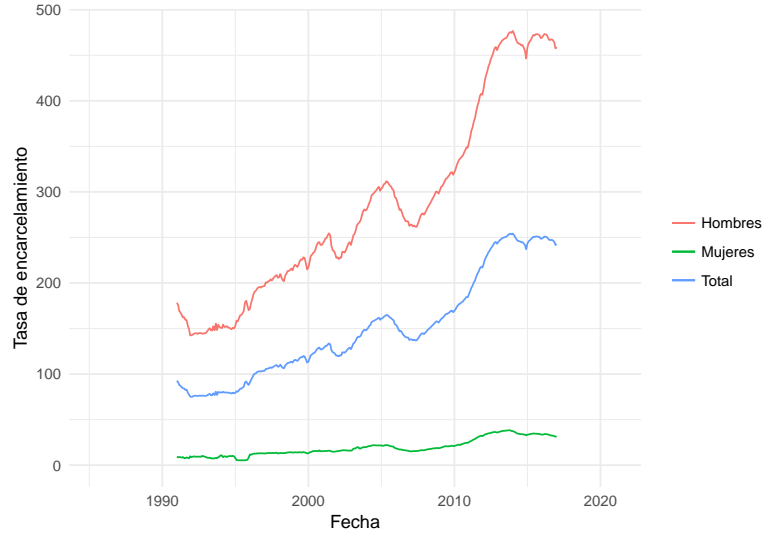


FIGURA 2.2. Tasa de encarcelamiento según género 1991 - 2017

Fuente: INPEC Elaboración propia

La tasa de encarcelamiento es un indicador que varía según la edad y género, siendo más elevado en los hombres que en las mujeres y más alto en los hombres jóvenes, que en los hombres mayores [1]. Otra posible explicación al cambio en la tasa de encarcelamiento es un cambio en la pirámide poblacional en el periodo analizado. No obstante, no podemos confirmar o refutar esta hipótesis pues la serie de tiempo, contenida en los datos de libre acceso no se encuentra desagregada por edad.

2.2. El sistema penitenciario en Colombia

La población carcelaria se ve afectada por dos políticas, la política penitenciaria, que determina las condiciones de privación de la libertad y la política criminal que determina las causas de encarcelamiento y la duración de las penas. [7]

A la población privada de la libertad antes del juicio se le denomina población sindicada, y a aquellos que han sido juzgados y se encuentran cumpliendo la sentencia se les denomina población condenada. La evolución de la población según situación jurídica se puede observar en la figura 2.3

2.2.1. Identificación del modelo

Podemos modelar el sistema penitenciario de la siguiente manera:

$$P_t = S_{t-1} + \alpha N_t - \gamma S_{t-1} \quad (2.1)$$

$$C_t = C_{t-1} - \omega C_{t-1} + \beta \gamma S_{t-1} \quad (2.2)$$

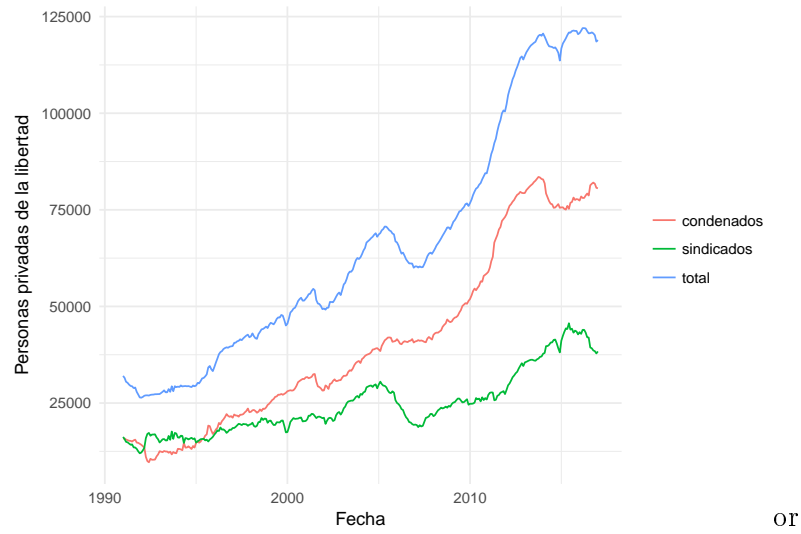


FIGURA 2.3. Población carcelaria por situación jurídica

Fuente: INPEC Elaboración propia

N_t = población nacional en el periodo t

S_t = población de sindicados en el periodo t

C_t = población de condenados en el periodo t

α = proporción de la población libre que ingresa al sistema carcelario

γ = proporción de sindicados que es juzgada cada periodo

β = proporción de sindicados que han sido encontrados culpables durante el juicio

ω = proporción de condenados que cumplen su pena cada periodo.

Modelos SARIMA

La estimación de los modelos se realiza usando el paquete base del software R [17], y el paquete *astsa* [20], cuyo uso es discutido en detalle por Shumway [19].

En adelante nos referiremos a la función de autocorrelación como ACF y a la función de autocorrelación parcial como PACF.

3.1. Identificación del modelo

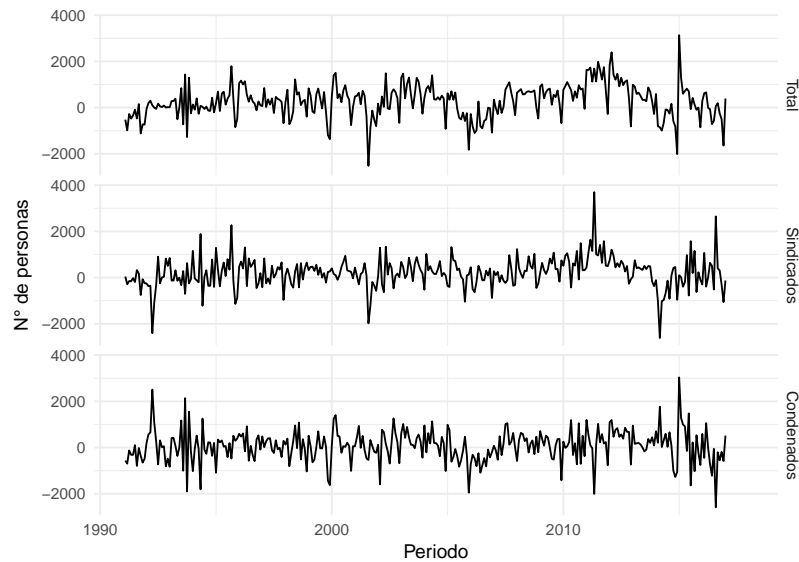


FIGURA 3.1. Variación inter-mensual de población carcelaria, sindicados y condenados

Fuente: INPEC Elaboración propia

Antes de realizar la proyección de una serie de tiempo, es necesario identificar el modelo que explique adecuadamente su comportamiento.

Aunque se conoce, por la estructura del proceso que genera los datos, que las series de población sindicada y condenada no son independientes, podemos simplificar la pro-

yección, tratándolas como independientes. En este caso, los modelos ARIMA y SARIMA resultan apropiados, pues permiten explicar separadamente cada observación en función del comportamiento histórico de la serie.

El capítulo anterior sugería que la población carcelaria, tanto sindicada como condenada, tiene una marcada tendencia al alza. En este caso una herramienta útil es mostrar gráficamente la variación mes a mes de la población. 3.1. Para tener una mejor estructura de análisis usamos la función "decompose" de R base.

La serie tiene un componente estacional marcado, con una reducción de la población carcelaria en diciembre. La variabilidad del componente aleatorio es elevada. La tendencia parece tener cambios estructurales en algunos periodos, por ejemplo reducción de la población carcelaria entre 2005-2007, y 2012 - 2015, e incrementos de la población de magnitud mayor al promedio entre 2008 y 2012.

La mayor parte de la variación de la población total se puede asociar con variaciones en a población sindicada 3.3. La población condenada tiene una tendencia con una tendencia más estable. 3.4

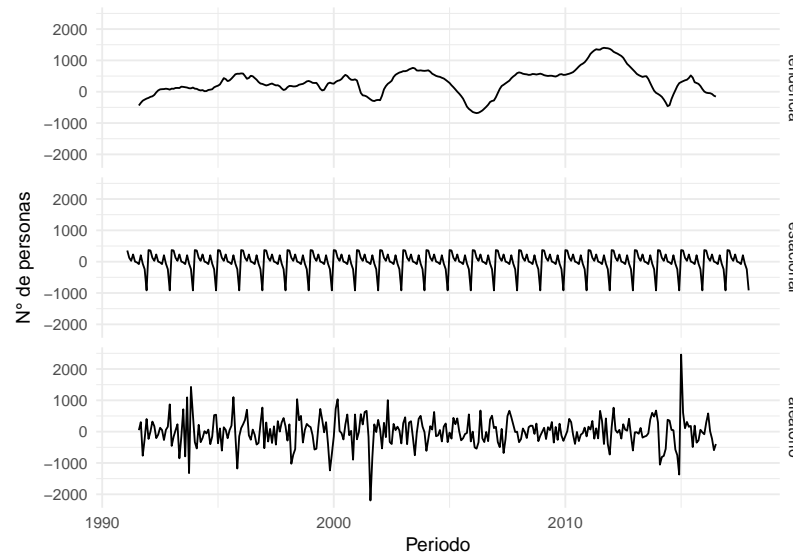


FIGURA 3.2. Variación mensual de población carcelaria descompuesta por tendencia, estacionalidad y componente aleatorio.

Fuente: INPEC Elaboración propia

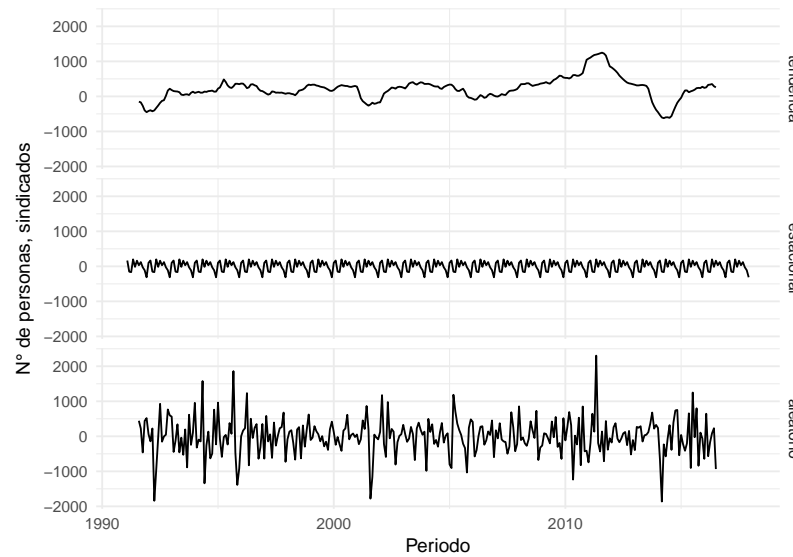


FIGURA 3.3. Variación mensual de población carcelaria sindicada, descompuesta por tendencia, estacionalidad y componente aleatorio.

Fuente: INPEC Elaboración propia

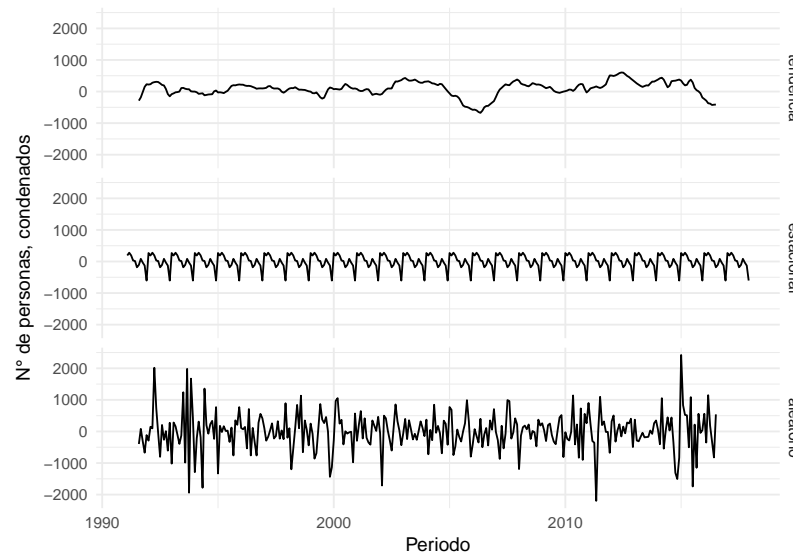


FIGURA 3.4. Variación mensual de población carcelaria condenada, descompuesta por tendencia, estacionalidad y componente aleatorio.

Fuente: INPEC Elaboración propia

Se trabaja sobre la diferencia de cada serie de tiempo con lag 1 y lag 12, al observar que la serie tiene un crecimiento sostenido entre 1991 y 2017, sugiriendo un proceso estocástico integrado. Sobre cada serie (población total, población sindicada y población condenada) se realiza la función de autocorrelación y autocorrelación parcial y se presenta en la figura 3.5. Estas funciones son usadas como herramienta de diagnóstico, para identificar modelos adecuados en cada serie.

Con base en la tabla 3.1 se realiza una revisión del comportamiento de las series de población carcelaria, población sindicada y población condenada.

TABLA 3.1. Comportamiento de la ACF y PACF en modelos ARMA(p,q) [19]

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

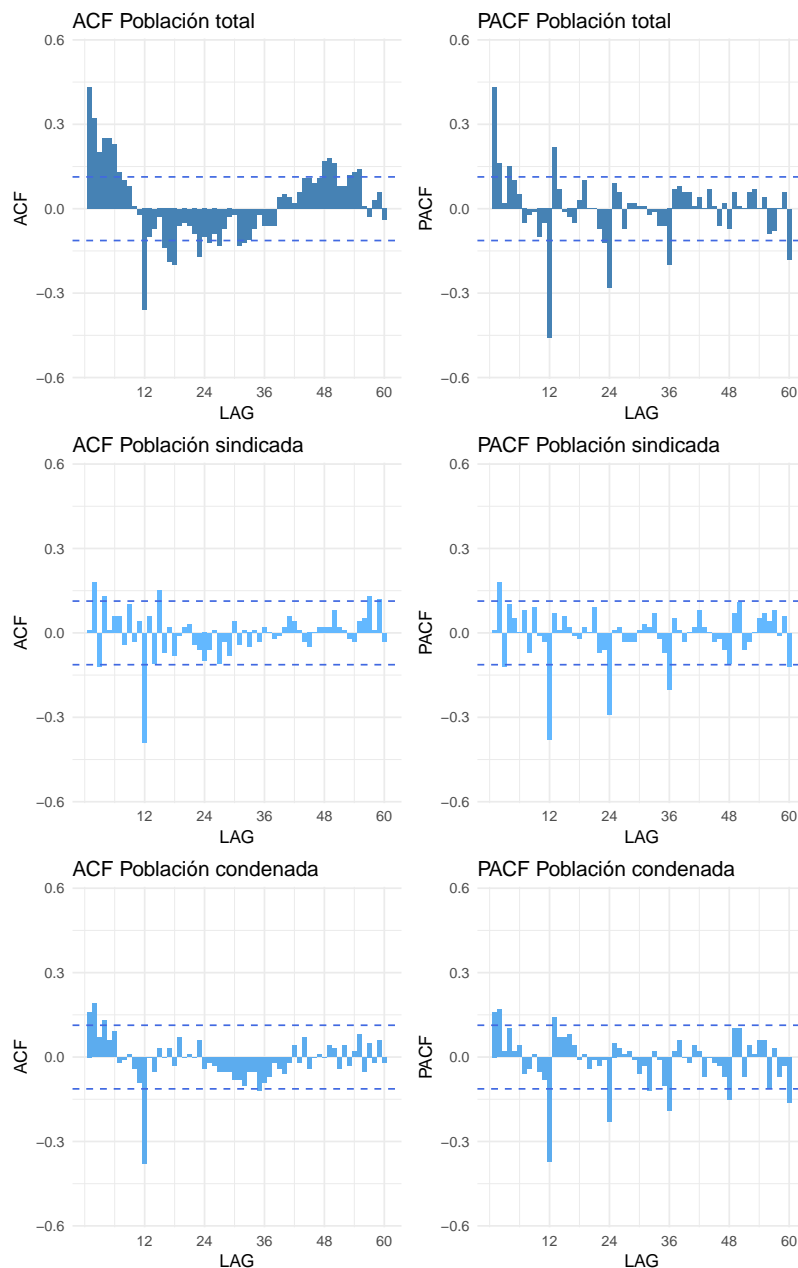


FIGURA 3.5. Autocorrelación parcial de la variación inter-mensual de la población

Fuente: INPEC Elaboración propia

La tabla 3.1 presenta una guía para la interpretación de la ACF y la PACF.

Si se analiza el total de la población carcelaria como una sola serie, es posible realizar una proyección separada, contra la cual contrastar las proyecciones de la población sindicada y de la condenada. En la figura 3.5 podemos observar que tanto la ACF como la PACF

decaen lentamente en la población total. La APCF decae progresivamente en los meses 12, 24, 36, mientras que la ACF se corta en el 24, lo que sugiere que el proceso es un promedio móvil de orden 1 en el componente estacional.

En este caso una primera estimación se realiza para la proyección de la población total como SARIMA(1,1,1,0,0,1).

En la figura 3.5 podemos observar que en la población sindicada la ACF decae lentamente, mientras la PACF cae por debajo del error luego del lag 1. La APCF decae progresivamente en los meses 12, 24, 36, mientras que la ACF se corta en el lag 12, lo que sugiere que el proceso es un promedio móvil de orden 1 en el componente estacional.

En este caso un primer acercamiento a la proyección de la población sindicada se presenta como SARIMA(1,1,1,0,0,1).

El comportamiento de la ACF y la PACF de la población condenada es similar al de la población sindicada, sugiriendo el mismo modelo: SARIMA(1,1,1,0,0,1).

3.2. Estimación de parámetros

Para cada serie (población total, población sindicada, población condenada) se evalúa el modelo sugerido por el análisis de la ACF y de la PACF; a partir de este modelo se propone un segundo modelo con un parámetro AR adicional, para identificar si se logra un ajuste mejor (medido con el AIC y el BIC); como tercer paso se utiliza la función *auto.arima* para identificar el modelo con menor AIC [9]; a partir del paso anterior se propone un modelo con menos parámetros, para buscar una solución más parsimoniosa.

De acuerdo con el diagnóstico se evaluaron los siguientes modelos para proyectar la población total:

1. Sarima de orden (1,1,1) y estacionalidad (0,0,2)
2. Sarima de orden (1,1,1) y estacionalidad (0,0,1). Se valida que el BIC y el AIC son menores en el modelo 1, como se esperaría de acuerdo con nuestro diagnóstico y se presenta en la tabla 3.2
3. Sarima de orden (3,2,1) y estacionalidad (0,0,2). Es el modelo sugerido por la función *auto.arima*, sin embargo incluye parámetros autoregresivos no significativos ($pvalue > 0.05$). El detalle de los resultados se presenta en la tabla 3.3
4. Sarima de orden (1,2,1) y estacionalidad (0,0,2). Este modelo tiene un BIC menor al modelo 3, y eliminando dos parámetros no significativos conserva un AIC cercano.

Para el modelo 1 y el modelo 2, Ljung-box sugiere que aún hay correlación significativa en los residuos, el modelo 3 y el modelo 4 no tienen correlación significativa en los residuos.

En las proyecciones siguientes para la población total se usará el modelo 4.

TABLA 3.2. Comparación de criterios de información

	Orden	AIC	BIC
1	(1,1,1)(0,0,2)	4885.2	4907.66
2	(1,1,1)(0,0,1)	4897.79	4916.51
3	(3,2,1)(0,0,2)	4874.46	4900.64
4	(1,2,1)((0,0,2)	4874.86	4893.56

TABLA 3.3. Parámetros del modelo (3,2,1,0,0,2)

	Estimate	SE	t.value	p.value
ar1	0.05	0.10	0.50	0.62
ar2	-0.09	0.07	-1.16	0.25
ar3	-0.16	0.07	-2.18	0.03
ma1	-0.68	0.09	-7.25	0.00
sma1	0.24	0.06	4.10	0.00
sma2	0.19	0.05	3.88	0.00

TABLA 3.4. Parámetros del modelo (1,2,1,0,0,2)

	Estimate	SE	t.value	p.value
ar1	0.19	0.08	2.39	0.02
ma1	-0.82	0.05	-16.48	0.00
sma1	0.24	0.06	4.10	0.00
sma2	0.19	0.05	3.89	0.00

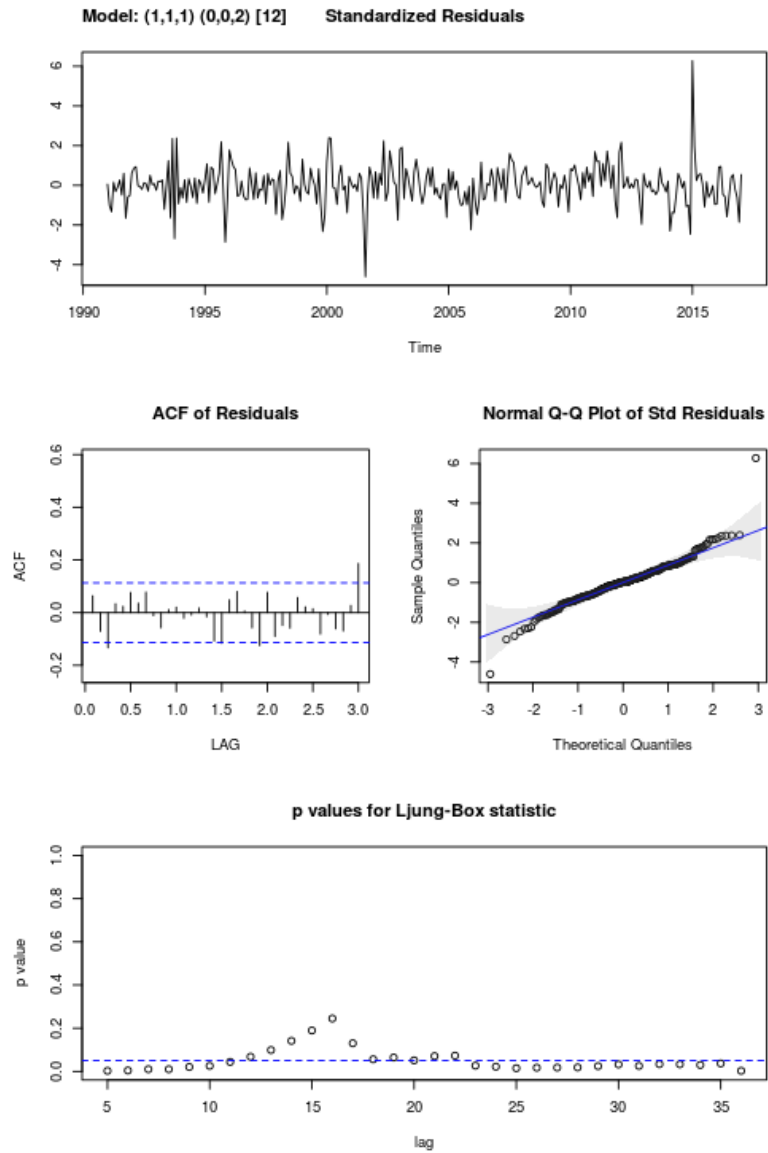


FIGURA 3.6. Diagnóstico del Modelo 1. sarima (1,1,1)(0,0,2), para la población carcelaria total.

Fuente: INPEC Elaboración propia

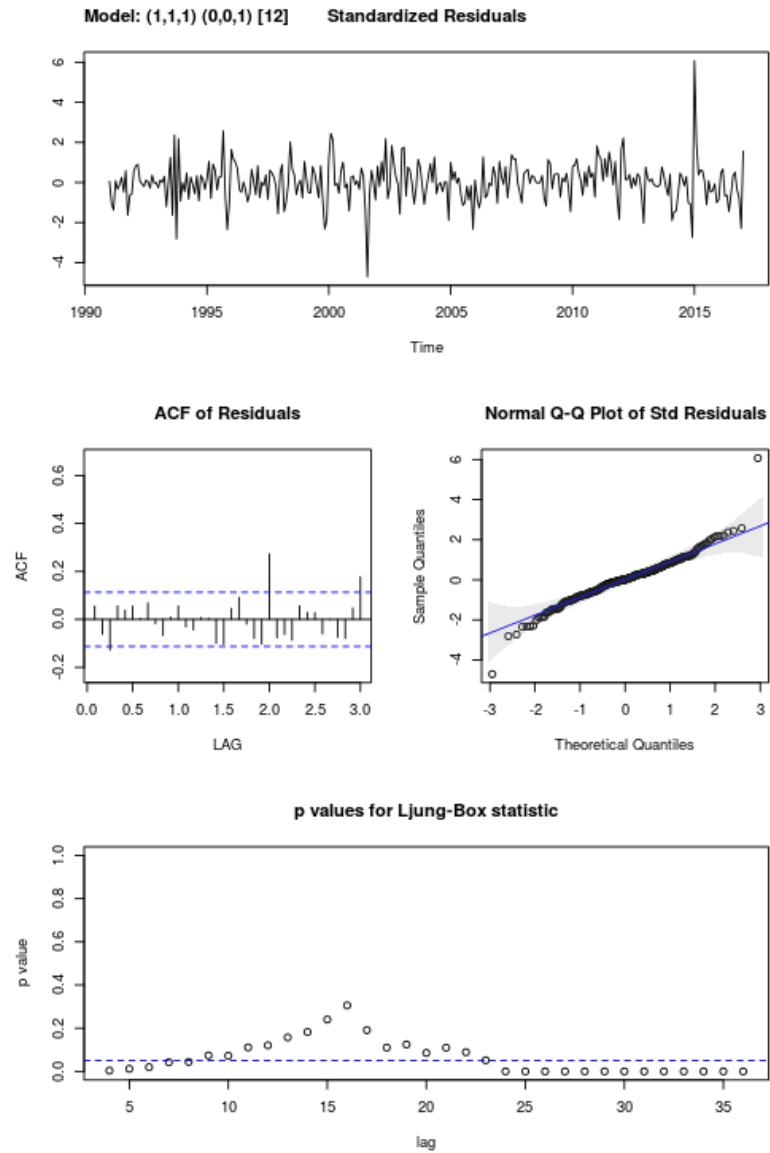


FIGURA 3.7. Diagnóstico del Modelo 2. sarima (1,1,1)(0,0,1), para la población carcelaria total.

Fuente: INPEC Elaboración propia

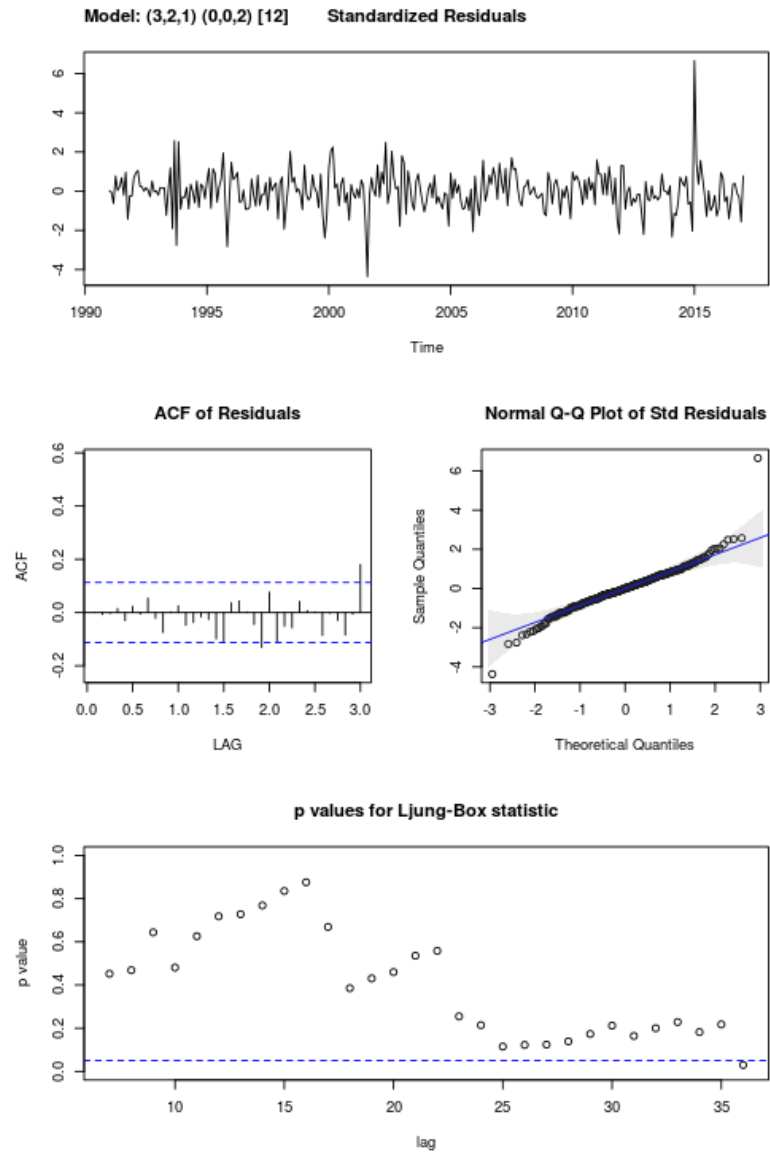


FIGURA 3.8. Diagnóstico del Modelo 3. sarima (3,2,1)(0,0,2), para la población carcelaria total.

Fuente: INPEC Elaboración propia

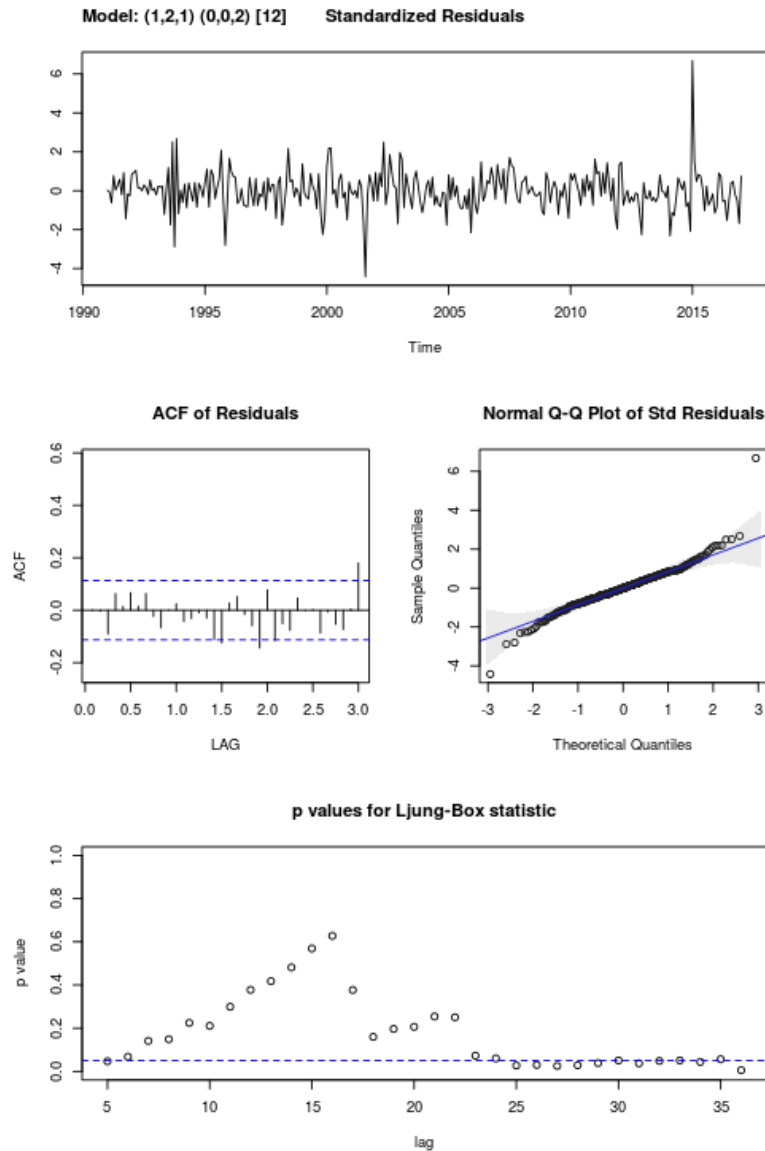


FIGURA 3.9. Diagnóstico del Modelo 4. sarima (1,2,1)(0,0,2), para la población carcelaria total.

Fuente: INPEC Elaboración propia

Para la población sindicada se usaron los siguiente modelos, de acuerdo con el diagnóstico realizado de la ACF y la ACPF.

1. Sarima de orden (1,1,1) y estacionalidad (0,0,1)
2. Sarima de orden (2,1,1) y estacionalidad (0,0,1). Se valida que el BIC y el AIC son menores que en el modelo 1, tiene el menor AIC de los modelos probados. 3.2
3. Sarima de orden (0,1,2) y estacionalidad (0,0,1). Es el modelo sugerido por la función auto.arima, tiene el menor BIC y un AIC cercano al modelo 2.
4. Sarima de orden (0,1,1) y estacionalidad (0,0,0). Este modelo tiene un BIC mayor al modelo 3, y a los dos modelos iniciales.

TABLA 3.5. Comparación de criterios de información

	ORDEN	AIC	BIC
1	(1,1,1,0,0,1)	4968.8	4987.52
2	(2,1,1,0,0,1)	4964.1	4986.56
3	(0,1,2,0,0,1)	4964.26	4982.98
4	(0,1,1,0,0,0)	4974.16	4985.39

Modelo 3 tiene el menor BIC y tiene pvalues for Ljung-Box que permiten descartar correlación. Los parámetros del modelo se presentan en la tabla 3.6 . El detalle del diagnóstico de cada modelo se encuentra en las figuras 3.10 a 3.13

TABLA 3.6. Parámetros del modelo (0,1,1,0,0,1)

	Estimate	SE	t.value	p.value
ma1	0.06	0.06	0.99	0.32
ma2	0.16	0.05	2.91	0.00
sma1	0.15	0.06	2.57	0.01
constant	64.94	53.21	1.22	0.22

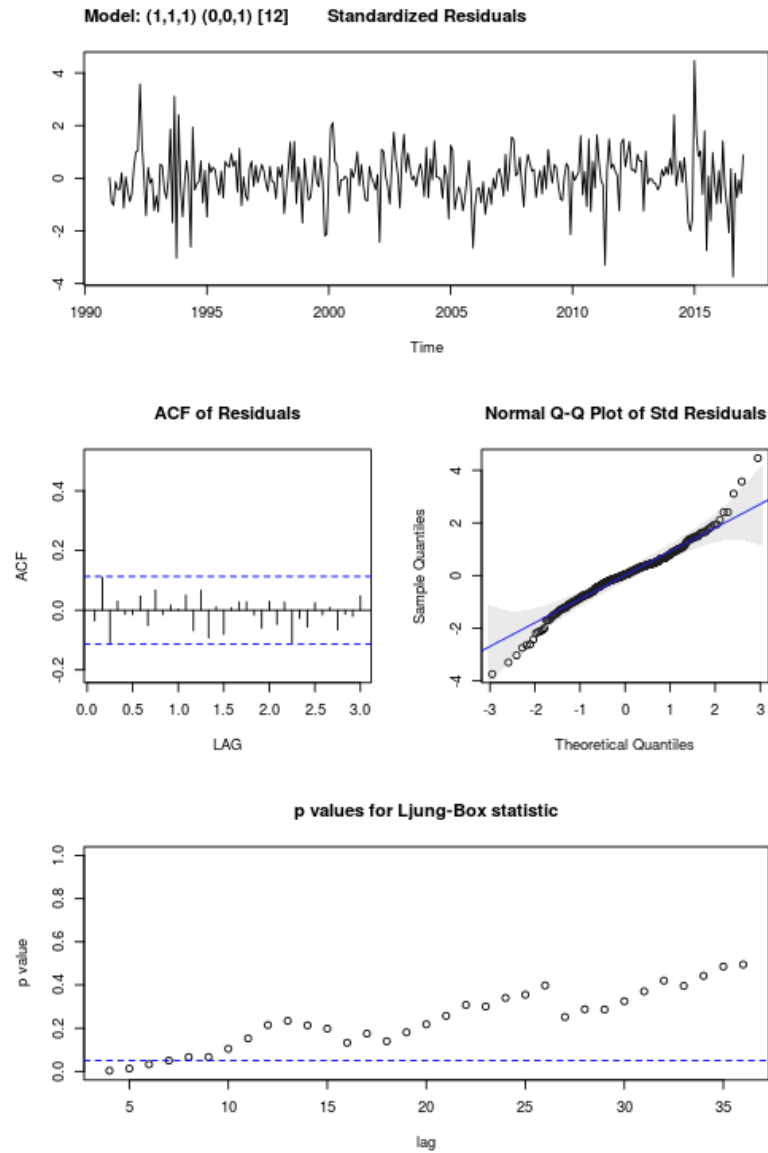


FIGURA 3.10. Diagnóstico del Modelo 1. sarima (1,1,1)(0,0,1), población carcelaria sindicada.

Fuente: INPEC Elaboración propia

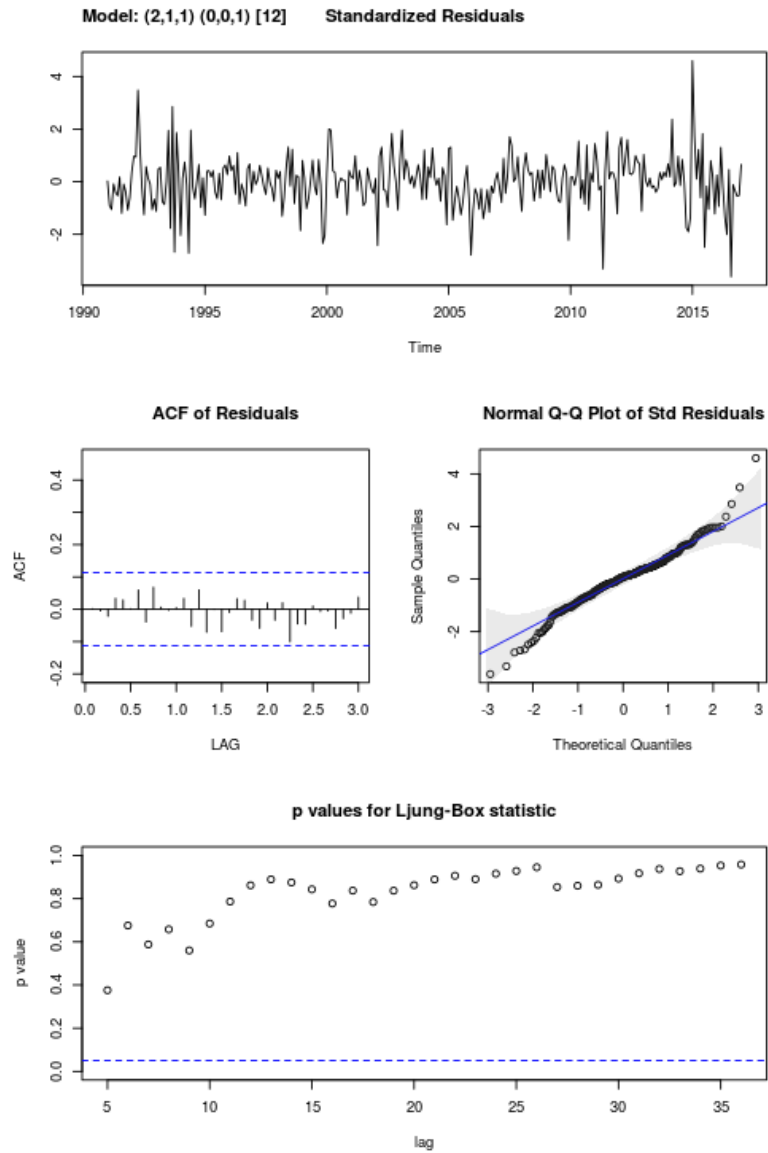


FIGURA 3.11. Diagnóstico del Modelo 2. sarima (2,1,1)(0,0,1), población carcelaria sindicada.

Fuente: INPEC Elaboración propia

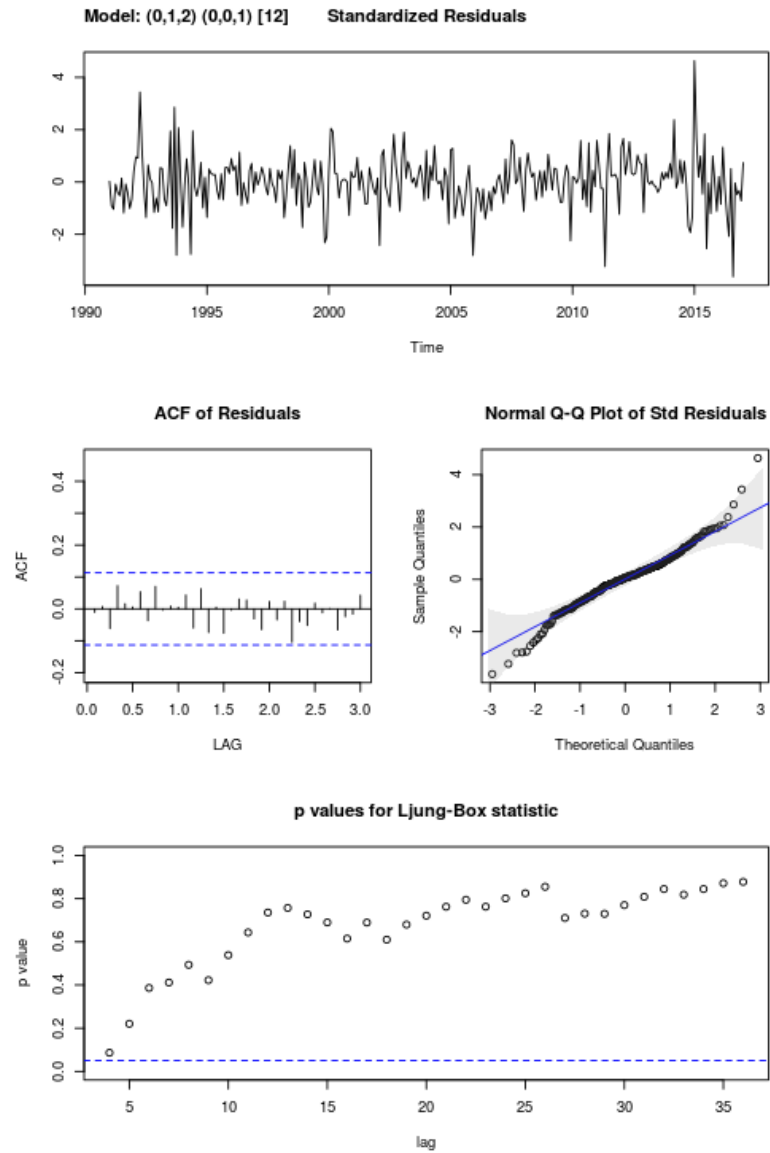


FIGURA 3.12. Diagnóstico del Modelo 3. sarima (0,1,2)(0,0,1), población carcelaria sindicada.

Fuente: INPEC Elaboración propia

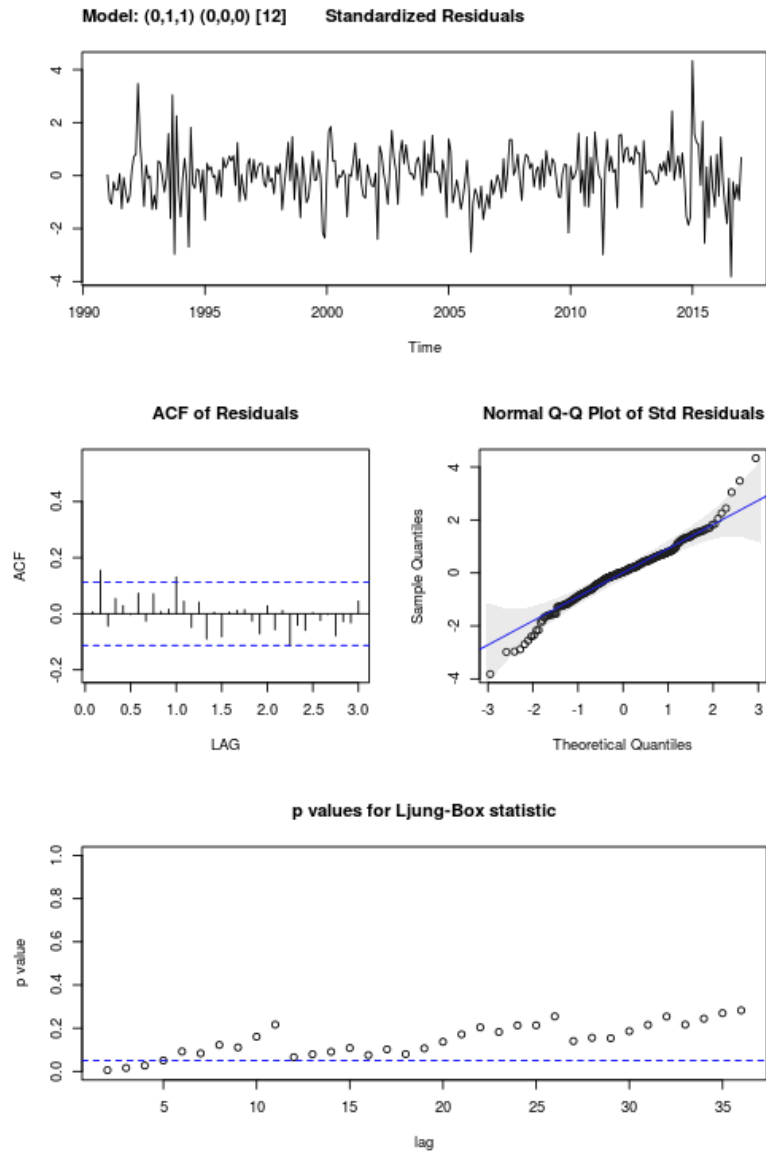


FIGURA 3.13. Diagnóstico del Modelo 4. sarima (0,1,1)(0,0,0), población carcelaria sindicada.

Fuente: INPEC Elaboración propia

Para la población condenada se utilizaron los siguientes modelos:

1. Sarima de orden (1,1,1) y estacionalidad (0,0,1).
2. Sarima de orden (2,1,1) y estacionalidad (0,0,1).
3. Sarima de orden (0,1,2) y estacionalidad (0,0,1).
4. Sarima de orden (0,1,1) y estacionalidad (0,0,0).

TABLA 3.7. Comparación de criterios de información

	ORDEN	AIC	BIC
1	(1,1,1,0,0,1)	4904.08	4922.79
2	(2,1,1,0,0,1)	4906	4928.46
3	(0,1,2,0,0,1)	4914.84	4933.55
4	(0,1,1,0,0,0)	4925.62	4936.85

Se escoge el modelo 1, pues aunque tiene menos parsimonia, además los pvalue de la estadística de Ljung-Box soportan rechazar la hipótesis de dependencia.

Los modelos 3 y 4 se utilizan para comparar con el comportamiento de la población sindicada. En este caso, el comportamiento de la población condenada es diferente y es bien definido por el modelo 1 con los parámetros presentados en la tabla 3.8.

TABLA 3.8. Parámetros del modelo (1,1,1,0,0,1)

	Estimate	SE	t.value	p.value
ar1	0.88	0.06	13.75	0.00
ma1	-0.73	0.09	-7.80	0.00
sma1	0.09	0.06	1.55	0.12
constant	198.62	86.28	2.30	0.02

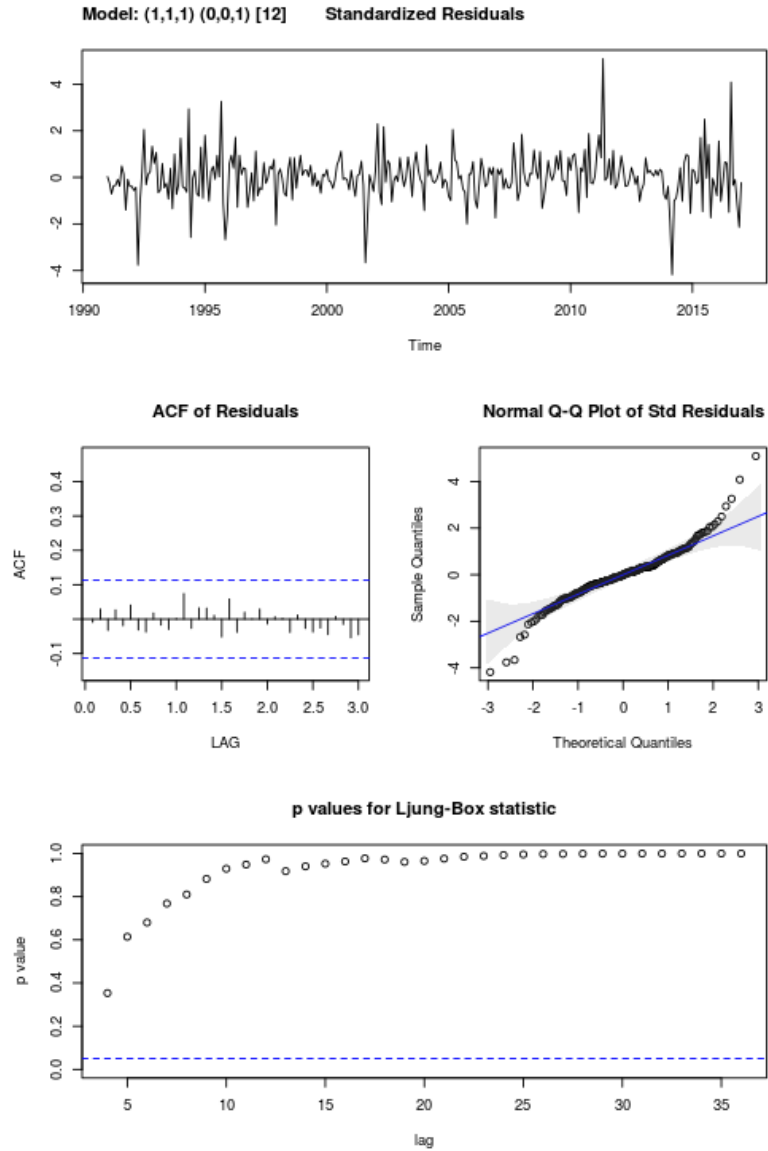


FIGURA 3.14. Diagnóstico del Modelo 1. sarima (1,1,1)(0,0,1), población carcelaria condenada.

Fuente: INPEC Elaboración propia

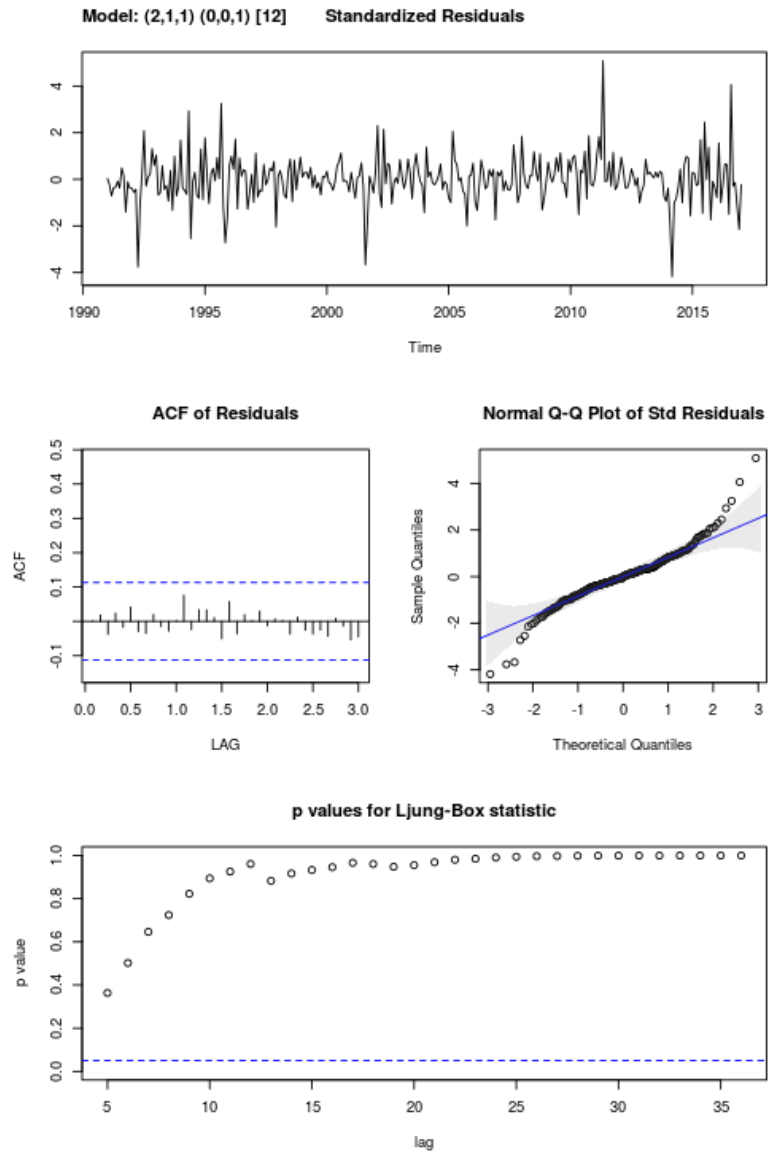


FIGURA 3.15. Diagnóstico del Modelo 2. sarima (1,1,1)(0,0,1), población carcelaria condenada.

Fuente: INPEC Elaboración propia

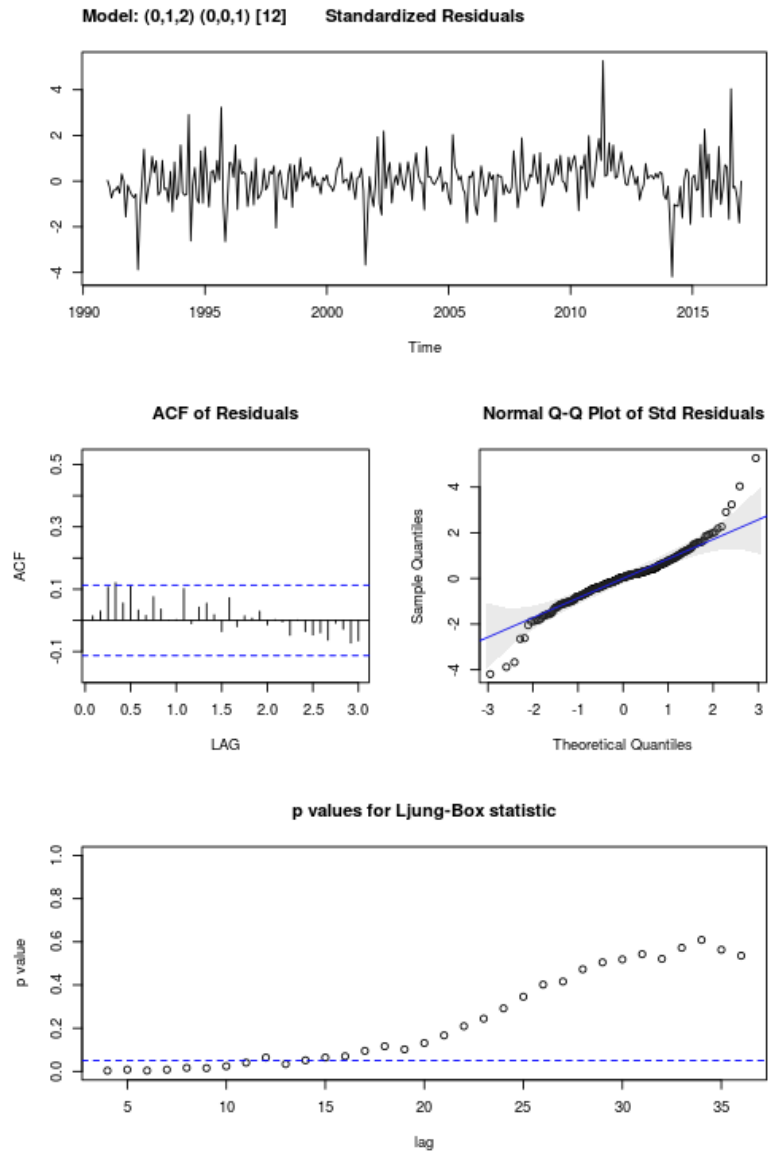


FIGURA 3.16. Diagnóstico del Modelo 3. sarima (0,1,2)(0,0,1), población carcelaria condenada.

Fuente: INPEC Elaboración propia

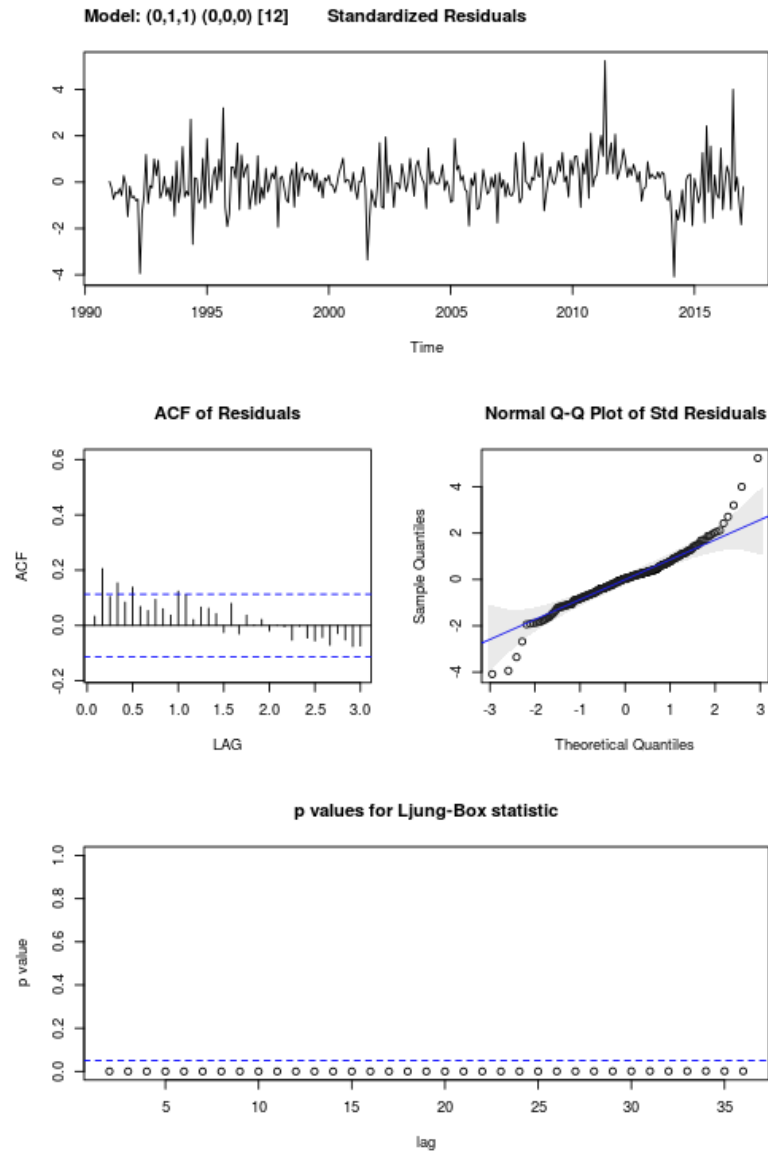


FIGURA 3.17. Diagnóstico del Modelo 4. sarima (0,1,1)(0,0,0), la población carcelaria condenada.

Fuente: INPEC Elaboración propia

3.3. Proyección

Proyección de la población total se presenta en la gráfica 3.18

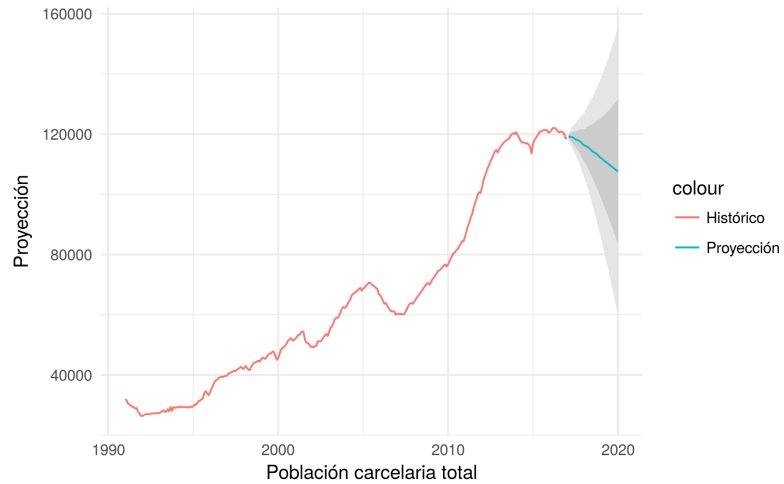


FIGURA 3.18. Proyección de la población carcelaria total, sarima (1,2,1,0,0,2)

Fuente: INPEC Elaboración propia

Proyección de la población sindicada se presenta en la gráfica 3.19

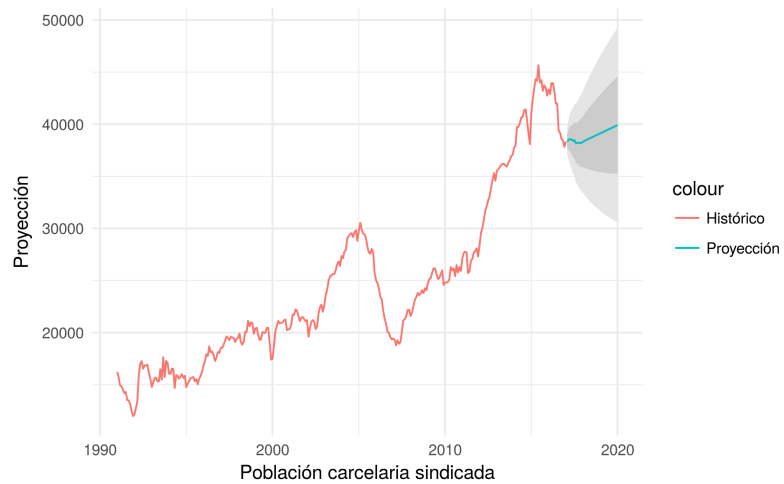


FIGURA 3.19. Proyección de la población carcelaria sindicada, sarima (0,1,1,0,0,1)

Fuente: INPEC Elaboración propia

Proyección de la población condenada se presenta en la gráfica 3.20

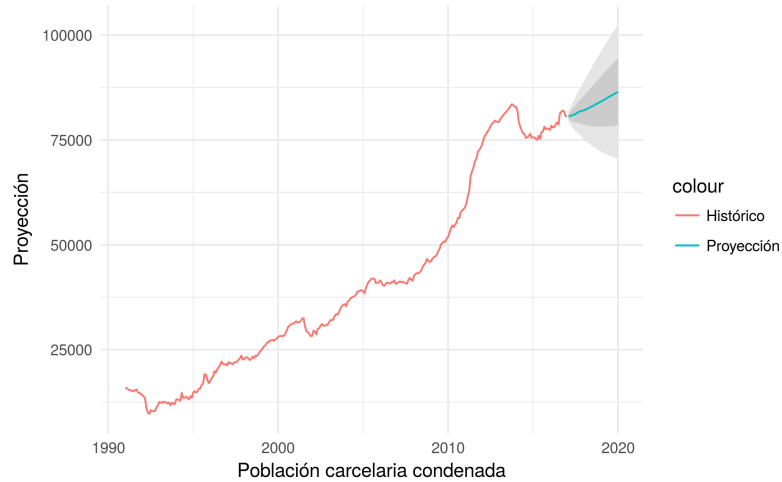


FIGURA 3.20. Proyección de la población carcelaria condenada, sarima (1,1,1,0,0,1)

Fuente: INPEC Elaboración propia

3.4. Conclusiones

Se seleccionó un modelo SARIMA para cada una de las series propuestas: población carcelaria total, población sindicada y población condenada. El residual en cada modelo seleccionado, ya no muestra correlación, y los estimadores de los parámetros son significativos. Con base en el modelo seleccionado para cada serie se realizaron proyecciones de población para 36 meses, la media de la proyección de la población total muestra una tendencia a la baja, mientras las proyecciones de la población sindicada y condenada tienen tendencia al alza.

Puesto que la población carcelaria total esta compuesta en su totalidad por la población sindicada y la población condenada, es de esperar que la proyección de la población total coincida en sentido y magnitud con la suma de las proyecciones de menor escala; sin embargo, la proyección de la población total difiere en sentido de la suma de las proyecciones.

Usar modelos ARIMA permite tener en cuenta la correlación del error y el comportamiento de las últimas observaciones, generando proyecciones precisas de corto plazo; sin embargo, dado que los modelos propuestos tienen un componente integrado, los intervalos de confianza se hacen significativamente más amplios, a medida que se alejan del inicio de la proyección, haciéndolos poco prácticos para proyecciones de largo plazo.

La elaboración de los intervalos de confianza descansa en el supuesto de normalidad del error, sin embargo, los Q-Q plot residuales muestran varias observaciones fuera de los intervalos de confianza para población sindicada y condenada, sugiriendo varios cambios de nivel o una distribución no normal del error. Contrastar los periodos donde el error está fuera de los intervalos de confianza contra los cambios en el sistema penitenciario, permitiría identificar cambios de nivel y controlarlos a través de variables dummy.

Situaciones como las encontradas en esta serie de tiempo (alta variabilidad, poblaciones compuestas por series de menor escala, procesos de migración inter-grupos, cambios de nivel asociados a cambios de régimen) pueden ser comunes en la proyección de poblaciones pequeñas. Estas situaciones, si se desean abordar usando las herramientas del análisis de series de tiempo pueden incluir: Modelos ARMAX, para los modelos cuya varianza

no es estable a través del tiempo, modelos autoregresivos de umbral (TAR) cuando se presenten cambios de régimen y vectores autoregresivos (VAR), para aquellas poblaciones cuyo comportamiento se relaciona con poblaciones pequeñas cercanas o con otros grupos de poblaciones especiales.

CAPÍTULO 4

Métodos demográficos para poblaciones pequeñas

No todas las personas tenemos la misma probabilidad de ser encarceladas. Los hombres tienen tasas de encarcelamiento más altas que las mujeres, y las personas entre 18 y 40 años son más susceptibles que los mayores de 40 [1]. Las cifras de encarcelamiento, presentadas en la figura 4.4 sugieren que estas tasas son estables en periodos cortos de tiempo y proporcionales entre los diferentes rangos de edad (si el encarcelamiento en un rango de edad se eleva sucede un incremento proporcional en los otros).

La población colombiana ha cambiado significativamente desde 1991, tanto en su magnitud como en su composición por edad. Hemos pasado de 35 millones en 1991 a 45 millones en la segunda década del siglo XXI. También hemos presenciado una reducción en la proporción de menores de 18 años, que ha pasado de XXX en 1991 a XXX en 2016. Se proyecta que en 2020 tendremos tantas personas entre 15 y 29 años, como entre 0 y 14, cuando en 1991 la diferencia era del XXX %. [6]

El grupo etario con mayor tasa de encarcelamiento está entre los 18 y 29 años. [1]. La población total en este grupo crece sostenidamente en el periodo 1991-2020. Como porcentaje de la población total, decrecen en dos puntos porcentuales en el mismo periodo, cediendo espacio a una población de mayor edad. El crecimiento de este grupo etario, particularmente en hombres podría ser un buen predictor de la población carcelaria.

Para incorporar la relación entre la población nacional y la población carcelaria se recurre al método censal-ratio, comúnmente usado en la proyección de poblaciones pequeñas. Este método requiere la inclusión de variables sintomáticas que agreguen información sobre la población pequeña.

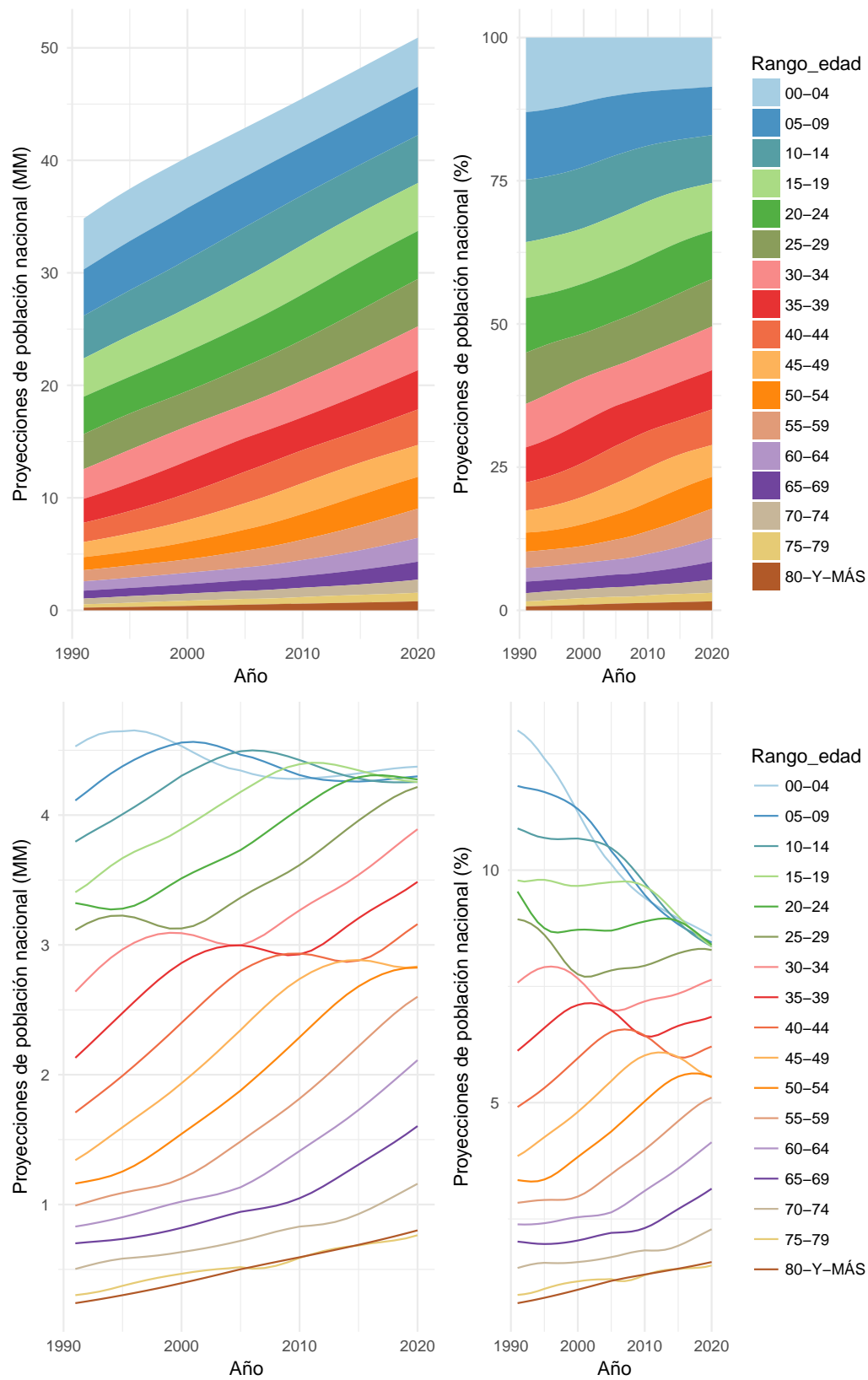


FIGURA 4.1. Proyección de la población nacional por rango de edad

Fuente: DANE Elaboración propia

4.1. Censal ratio

Dentro de los métodos para proyectar poblaciones pequeñas se encuentra censal-ratio (razón-censal). Éste método se vale una variable sintomática asociada a la población pequeña, a la que se supone es directamente proporcional. Un ejemplo de variable sintomática es la cantidad de nacimientos en una ciudad, con respecto a su población total; otro ejemplo es la cantidad de niños enrolados en un curso, con respecto a un rango etario en determinado espacio geográfico.

El censal-ratio supone que la relación entre la variable sintomática y la población a estimar es una variable aleatoria cuyo valor esperado no cambia a través del tiempo.

Cuando se cuenta con estimaciones independientes de la población total en periodos inter-censales se realiza una corrección sobre las proyecciones de cada área, tal que la suma de las proyecciones de áreas pequeñas sea igual a la estimada independientemente. La corrección mantiene la proporción entre proyecciones de áreas pequeñas.

Un uso frecuente de los métodos de proyección para poblaciones pequeñas es realizar interpolación en periodos no censales. Uno de los inconvenientes mencionados es la dificultad de realizar extrapolación, puesto que para los periodos post-censales, los indicadores no están necesariamente disponibles. Una posible solución es realizar la estimación usando periodos rezagados o utilizar la última proyección de la variable sintomática como proyección. [21].

Siguiendo a Swanson [21] tenemos:

$$R_{A,t,j} = \frac{S_{A,t,j}}{P_{A,t}}$$

R = Razón entre variable sintomática y población pequeña

A = A-ésima área pequeña

S = Valor de la variable sintomática

t = periodo

j = indicador

Si tenemos un R observado en el periodo T tendremos que la proyección del periodo $t + k$ será:

$$\hat{P}_{A,t+k} = \frac{S_{A,t+k,j}}{R_{A,T,j}}$$

Cuando se tiene un estimado independiente para la población padre se realiza un control, de manera que la suma de proyección para las áreas pequeñas sea igual a la proyección del área padre:

$$\hat{Q}_{A,t+k} = \frac{P_{A,t+k,j}}{\sum P_{A,t+k,j}} * P_{t+k} \quad (4.1)$$

$\hat{Q}_{A,t+k}$ = Proyección corregida para el área A

P_{t+k} = Proyección independiente total en el periodo $t + k$

La variable sintomática cumple la tarea de agregar información conocida al modelo para hacerlo más preciso. En un primer paso se considera la población privada de la libertad una población pequeña, dentro de la población nacional. Dentro de las estadísticas publicadas periódicamente por el INPEC no se han encontrado variables sintomáticas que cubran más de dos años [10], razón por la cual se realiza el siguiente proceso: se utiliza una constante en el lugar de la variable sintomática, lo que da lugar a una proyección constante por edad genero; luego ajustamos la proyección al total nacional por rango de edad proyectado por el DANE. Este modelo es equivalente a proyectar una tasa de encarcelamiento constante para cada rango de edad. Para el periodo 2017 - 2020 se utiliza la tasa de encarcelamiento.

Aunque no conocemos el comportamiento de la tasa de encarcelamiento específica por edad, conocemos la población carcelaria total y la tasa de encarcelamiento. Usamos el supuesto de proporción estable por rango etario, para ajustar la proyección 2005 - 2016 a la población carcelaria total.

$$R_e = \frac{1}{P_{c,e,2017}}$$

$$\hat{P}_{c,2017+k} = P_{c,e,2017}$$

y

$$\hat{Q}_{c,e,2017+k} = (\hat{P}_{c,e,2017} * P_{e,2017}) * P_{e,2017+k}$$

donde:

$\hat{Q}_{c,e,2017+k}$ = Proyección de población carcelaria ajustada, para el rango etario, en el periodo $2017 + k$

$P_{c,2017}$ = Población carcelaria en 2017, a mitad de periodo, en el rango etario e.

$P_{e,2017}$ = Población estimada 2017 DANE a mitad de periodo en el rango etario e.

$P_{e,2017+k}$ = Población estimada en $2017 + k$, del DANE a mitad de periodo en el rango etario e.

Al aplicar esta proyección a periodos anteriores a 2017 tenemos una estimación de lo que habría sido la población carcelaria, en un escenario de tasas de encarcelamiento específicas constantes. Al comparar esta proyección con la población privada de la libertad observada, es posible evidenciar el impacto que ha tenido el proceso de cambio demográfico en la población carcelaria, además de poner en evidencia que no es el principal factor detrás de este cambio. El incremento debe además corresponder con un incremento de las tasas de encarcelamiento específicas.

Para ajustar la proyección de las tasas de encarcelamiento específicas se trata la población carcelaria como un área padre y a las poblaciones por rango etario como poblaciones pequeñas. Usando la ecuación 4.1 tenemos:

$$\hat{V}_{c,e,2017+k} = (\hat{P}_{c,e,2017} * \hat{P}_{c,2017+k}) * P_{c,2017+k}$$

donde:

$\hat{V}_{c,e,2017+k}$ = Población carcelaria ajustada por edad.

$\hat{P}_{c,2017+k}$ = Población carcelaria estimada por genero.

$P_{c,2017+k}$ = Población carcelaria observada.

4.2. Población privada de la libertad por rango de edad Febrero 2016 a Febrero 2018

El INPEC publica un reporte mensual con estadísticas detalladas de la población carcelaria al cierre de mes. Los reportes están disponibles en la página del INPEC desde e incluyen la composición de la población privada de la libertad desde Enero 2013. La estructura de los reportes no es uniforme para todos los periodos. La composición por rangos etarios antes de febrero 2016 consolidaba en las categorías: 18 a 29 años, 29 a 54 años, 55 a 64 años, y mayor de 65 años. A partir de marzo de 2016 se agrupa 18 a 24 y grupos quinquenales hasta los 65 años. [10]

En la figura 4.2 podemos observar la composición demográfica de la población privada de la libertad en Junio 2017. Es una población compuesta mayoritariamente por hombres. Tanto en hombres como mujeres se observa un patrón similar en la estructura etaria, el punto más alto se encuentra entre 25 y 29 años, 18 a 25 y 29 a 35 tienen un orden de magnitud comparable. Después de los 35 años la población se reduce en cada grupo quinquenal.

En la figura 4.3 se presenta la población privada de la libertad, separada por genero y edad desde marzo 2016 hasta marzo 2018. El periodo de análisis se escogió a conveniencia, considerando la disponibilidad de datos del INPEC con grupos etarios uniformes. La figura 4.3 muestra que la población carcelaria mantiene la forma de campana, con pico en 25-29, en todos los meses analizados, tanto en hombres como mujeres. En hombres el rango de edad con mayor variabilidad es el rango 18-24. En mujeres la variabilidad se presnta en los rangos 18-24, 25-29 y 30-34. [18]

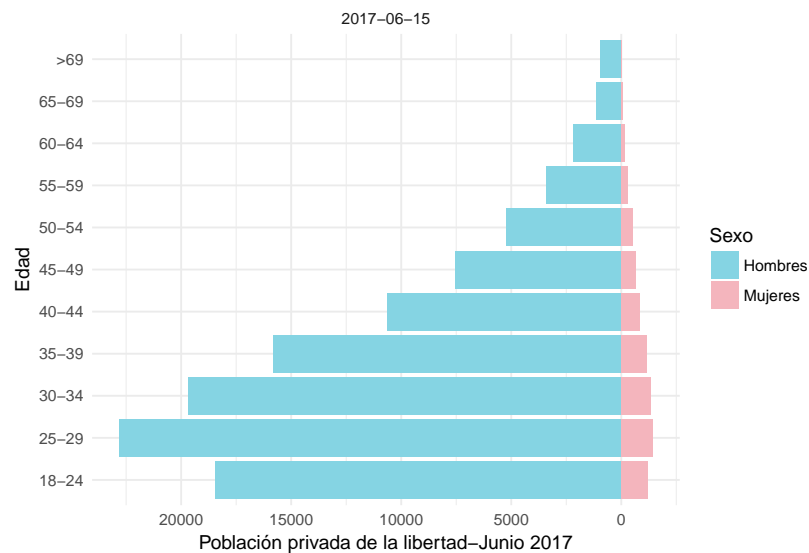


FIGURA 4.2. Piramide poblacional, población privada de la libertad, Junio - 2017

Fuente: INPEC Elaboración propia

Esta variabilidad podría estar asociada a un componente estacional o a un componente aleatorio. Aunque la longitud de la serie disponible no permite realizar inferencia sobre

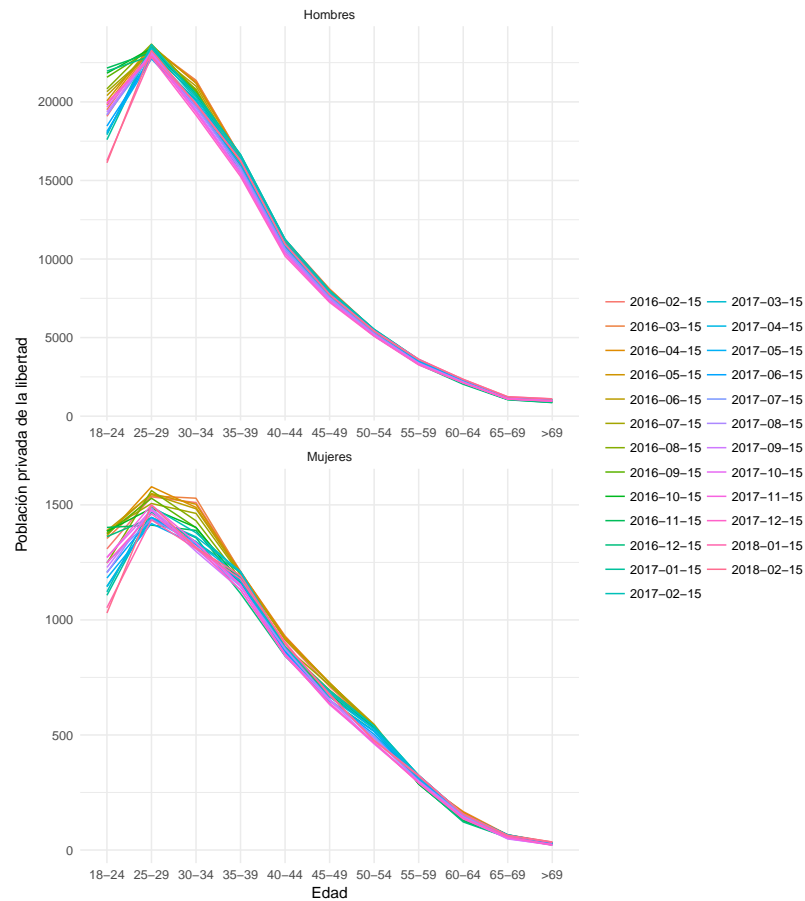


FIGURA 4.3. Población privada de la libertad. Febrero-2016 a Febrero-2018

Fuente: INPEC Elaboración propia

componentes de estacionalidad y tendencia, la figura muestra que: A través del año el grupo 18-24 crece, para caer abruptamente en enero del año siguiente; comportamiento repetido en enero 2016 y enero 2017. El comportamiento es similar en mujeres una vez se controla por el orden de magnitud. El comportamiento en las demás series es similar, se reduce durante todo el año, para subir nuevamente en enero. Este comportamiento es coherente en hombres y mujeres, una vez se controla la magnitud. 4.3 Este comportamiento implica que no es conveniente trabajar la proyección de la tasa de encarcelamiento a nivel mensual pues no se puede estimar el aparente componente de estacionalidad de la serie. Por conveniencia se trabaja con la proyección de población a mitad de periodo, y se trata la tasa de encarcelamiento como una variable aleatoria cuyo valor esperado se mantiene a través del tiempo.

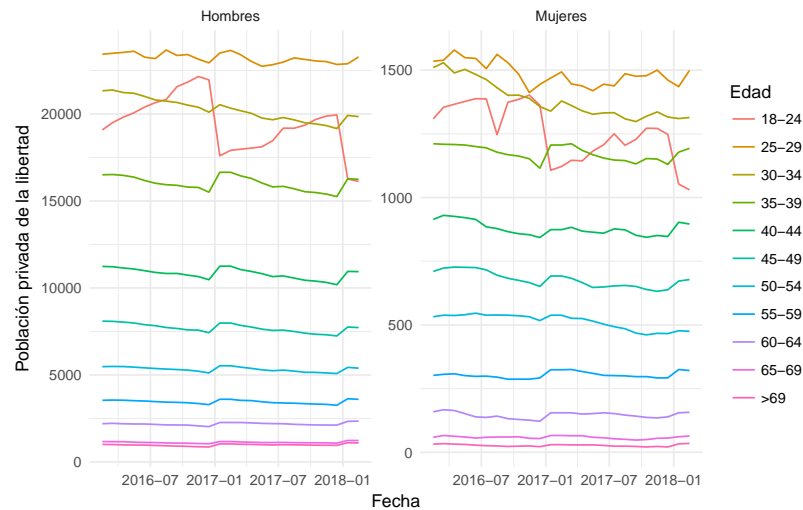


FIGURA 4.4. Población privada de la libertad por rango etario. Febrero-2016 a Febrero-2018

Fuente: INPEC Elaboración propia

Puesto que no se cuenta con información de periodos más extensos se valida análisis a un menor nivel de granularidad. Dentro de la información disponible se encuentra separada la población carcelaria por regional. En la figura 4.5 se observa nuevamente, en todas las regionales, la población carcelaria con forma de campana. Consistentemente las mujeres tienen mayor variabilidad que los hombres en cada rango de edad y poblaciones carcelarias más pequeñas se encuentran asociadas a una mayor variabilidad.

La forma del aparente componente estacional también se replica a nivel regional. Ver figura 4.6.

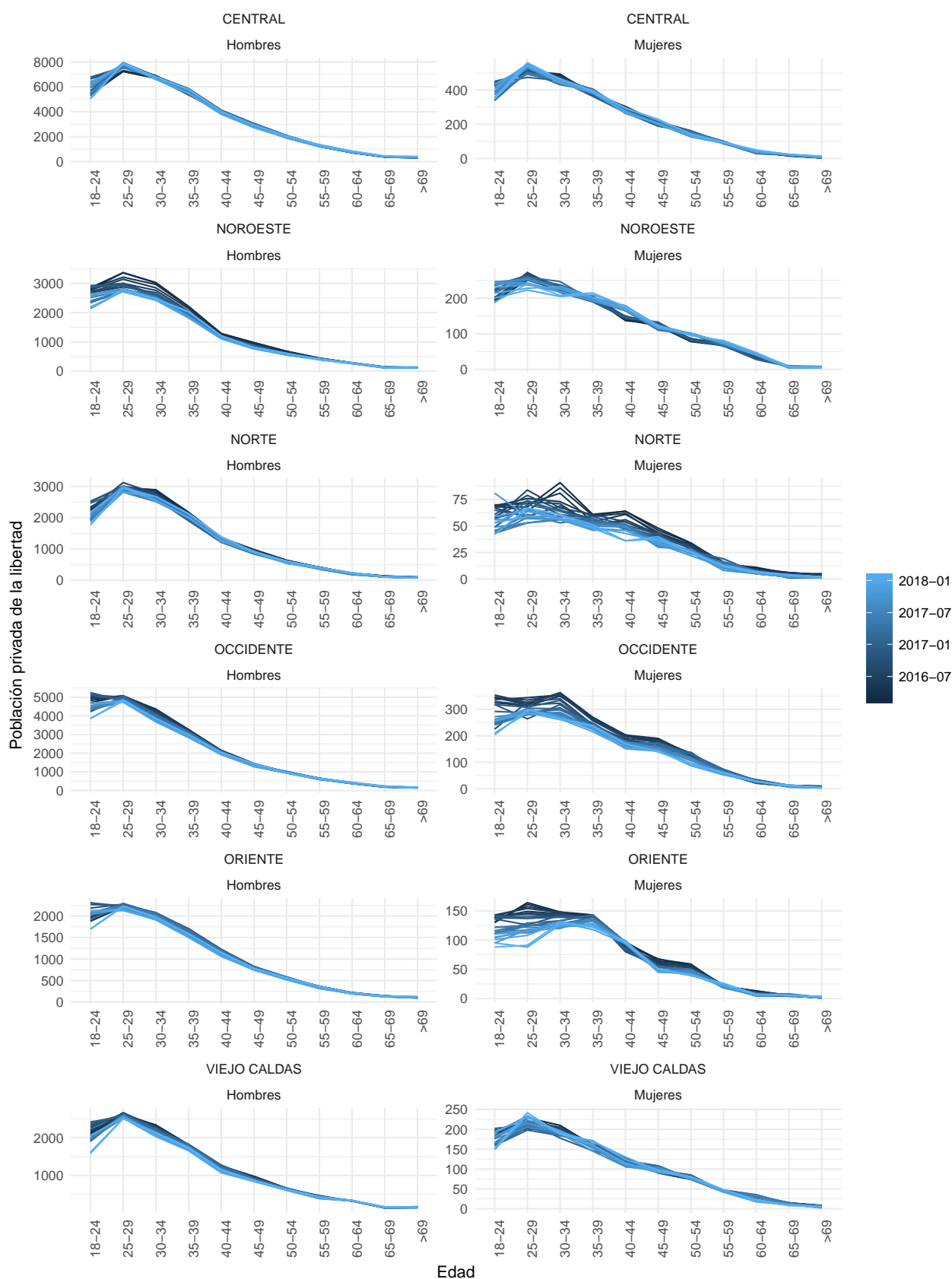


FIGURA 4.5. Población privada de la libertad por rango etario, región. Febrero-2016 a Febrero-2018

Fuente: INPEC Elaboración propia

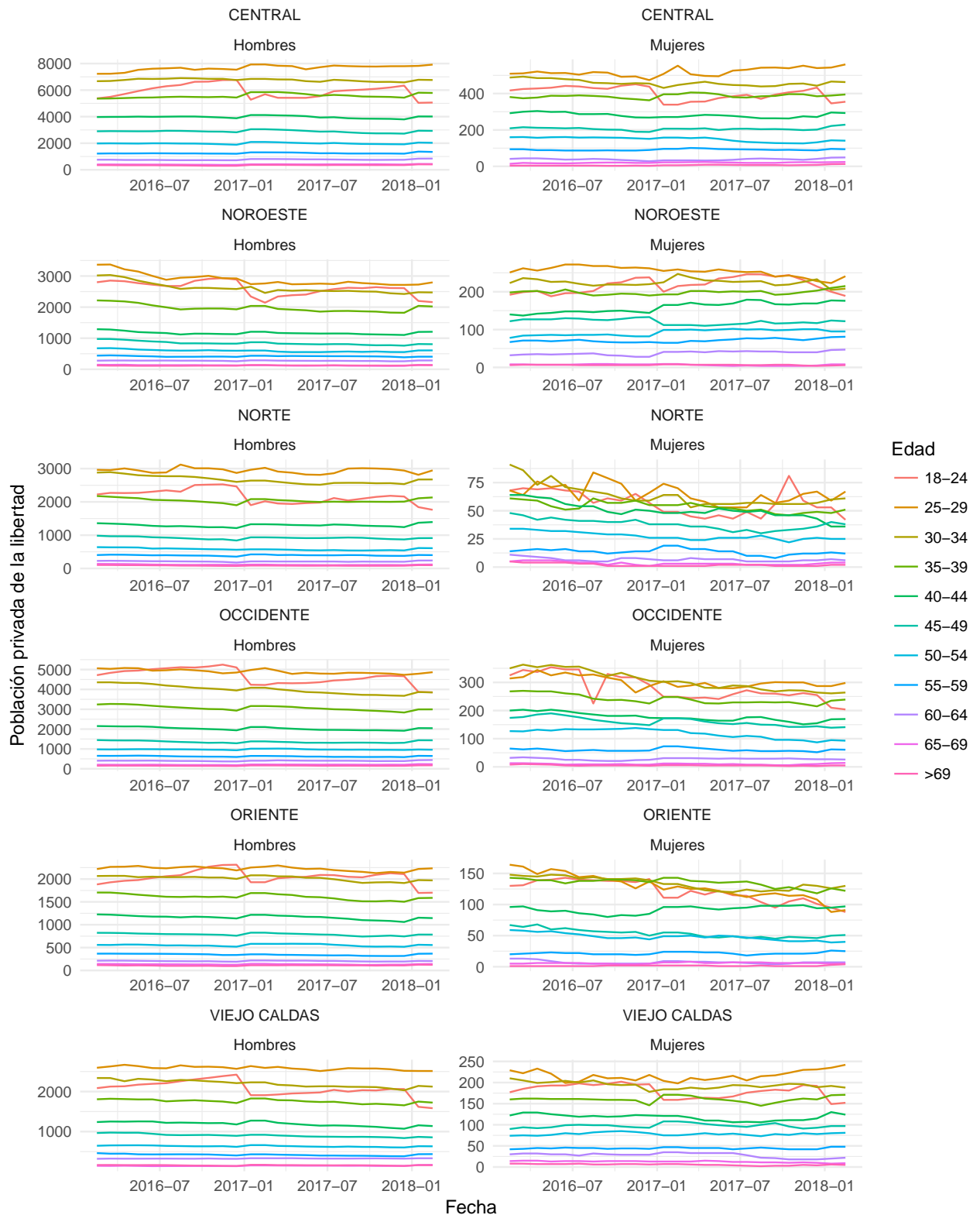


FIGURA 4.6. Población privada de la libertad por rango etario, región. Febrero-2016 a Febrero-2018

Fuente: INPEC Elaboración propia

Las personas no aparecen de la nada en la cárcel.

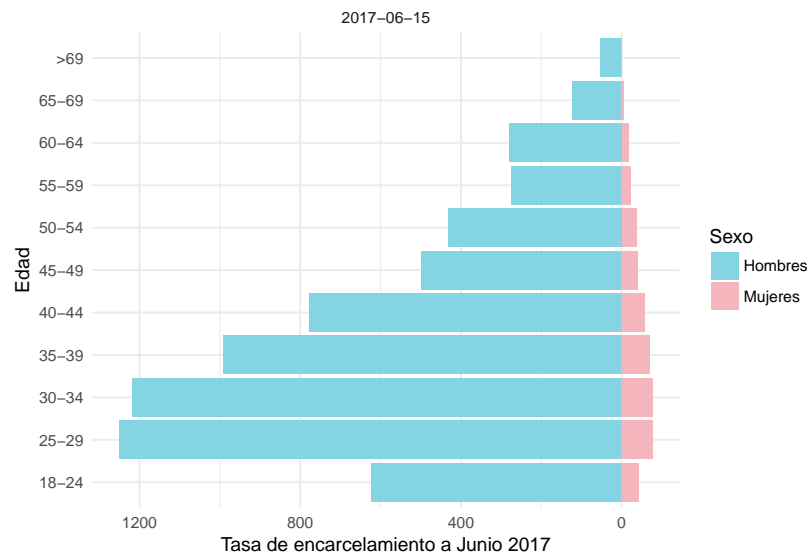


FIGURA 4.7. Tasa específica de encarcelamiento. Junio-2017

Fuente: INPEC, DANE Elaboración propia

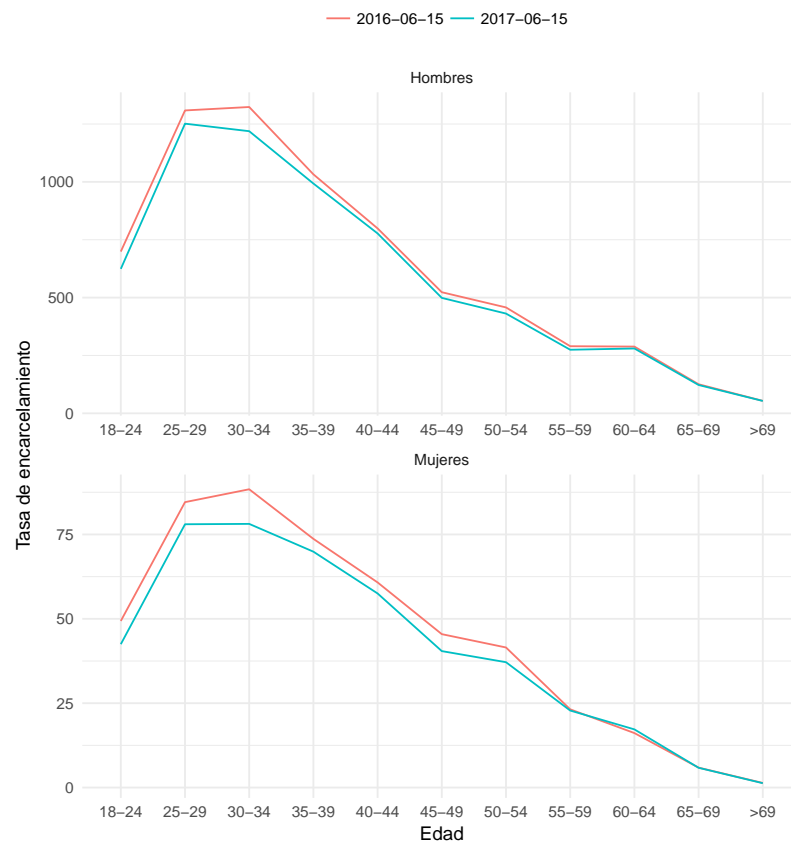


FIGURA 4.8. Tasa específica de encarcelamiento. Junio-2016, Junio -2017

Fuente: INPEC, DANE Elaboración propia

4.3. Proyecciones Censal Ratio

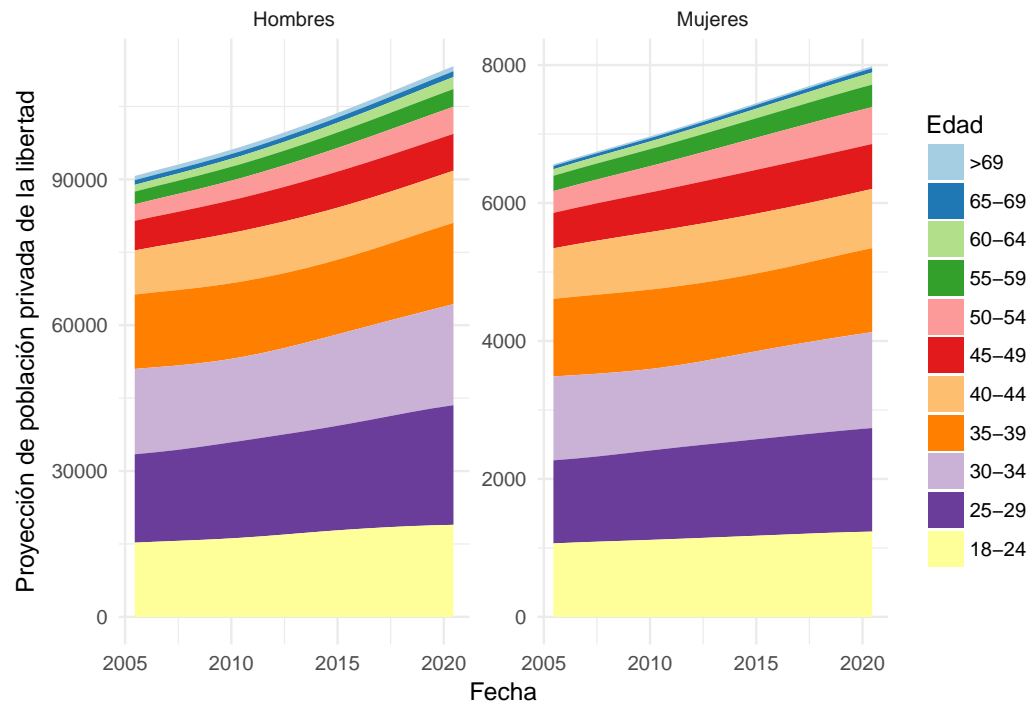


FIGURA 4.9. Proyección de la población privada de la libertad, 2005 - 2020, ajustado por rango etario DANE

Fuente: INPEC, DANE Elaboración propia

CAPÍTULO 5

Modelos estado-espacio

5.1. Selección de Software

Para mantener un único ambiente de desarrollo, nos enfocamos en los paquetes desarrollados en R, como los mencionados por Petris [16], Helske [8] y Tusell [22]. Particularmente usaremos el paquete *Kfas* por Helske(2017) [8]

5.2. Identificación del modelo

5.3. Simulación Monte-Carlo

5.4. Estimación de parámetros

5.5. Proyecciones 2017 - 2020

5.6. Conclusiones

Conclusiones

- La escritura de tesis utilizando L^AT_EX permite que se obtengan documentos de una presentación elegante, agradable, de una impresión incomparable, de escritura bastante simple en cuanto al texto técnico y de formulas matemáticas, junto con un manejo automático del formato de las partes de un documento y las referencias bibliográficas, desprendiéndose así de los detalles de edición que en otras herramientas, producen tantas frustraciones y dolores de cabeza.

Trabajo futuro

- Implementar y corregir todos aquellos errores que los usuarios de esta plantilla puedan encontrar, así como las sugerencias para la modificación de la plantilla que sean pertinentes.

Bibliografía

- [1] Alfred Blumstein, Jacqueline Cohen, and Harold D. Miller, *Demographically disaggregated projections of prison populations*, Journal of Criminal Justice **8** (1980), no. 1, 1–26.
- [2] George E P Box, Gm Jenkins, and Gc Reinsel, *Time series analysis: forecasting and control*, 2013.
- [3] Paolo Brandimarte, *Time Series Models*, Quantitative Methods: An Introduction for Business Management, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 527–579.
- [4] Peter J. Brockwell, *Time Series*, International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 1601–1605.
- [5] Jacques J.F Commandeur and Siem Jan Koopman, *An Introduction to State Space Time Series Analysis*, (2007), 189.
- [6] DANE, *Estimaciones población DANE*, 2009.
- [7] Departamento Nacional de Planeación, *Conpes 3828 POLÍTICA PENITENCIARIA Y CARCELARIA EN COLOMBIA*, (2015).
- [8] Jouni Helske, *Exponential Family State Space Models in R*, Journal of Statistical Software **78** (2017), no. 10, 1–39.
- [9] Rob J Hyndman, *{forecast}: Forecasting functions for time series and linear models*, 2017.
- [10] INPEC, *Estadísticas - Estadística - INPEC*, 2018.
- [11] Institute for Criminal Policy Research, *Colombia / World Prison Brief*, 2016.
- [12] Ministry Justice, *Prison Population Projections 2014 - 2020, England and Wales*, Home Office Statistical Bulletin (2014), no. November, 31.
- [13] Ronald Demos Lee and Shripad Tuljapurkar, *Stochastic Population Forecasts for the United States: Beyond High, Medium, and Low*, Journal of the American Statistical Association **89** (1994), no. 428, 1175.
- [14] Helmut Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

-
- [15] Todd D Minton, Bjs Statistician, Scott Ginder, Susan M Brumbaugh, Hope Smiley-Mcdonald, and Harley Rohloff, *Census of Jails: Population Changes, 1999 - 2013*, (2015).
 - [16] Giovanni Petris and Sonia Petrone, *State Space Models in R*, Journal of Statistical Software **41** (2011), no. 4, 128–129.
 - [17] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017.
 - [18] Karthik Ram and Hadley Wickham, *wesanderson: A Wes Anderson Palette Generator*, 2018.
 - [19] Robert H Shumway and David S Stoffer, *Time Series Analysis and Its Applications*, Springer Texts in Statistics, vol. 45, Springer New York, New York, NY, 2011.
 - [20] David Stoffer, *astsa: Applied Statistical Time Series Analysis. R package version 1.3*, 2014.
 - [21] David A Swanson and Jeff Tayman, *Subnational Population Estimates*, The Springer Series on Demographic Methods and Population Analysis, vol. 31, Springer Netherlands, Dordrecht, 2012.
 - [22] Fernando Tusell, *Kalman Filtering in R*, Journal of Statistical Software **39** (2011), no. 2, 1–27.
 - [23] Wai-Yin Wan¹, Steve Moffatt², Zachary Xie³, Simon Corben, and Don Weatherburn, *Forecastin prison populations using sentencing and arrest data*, Tech. report.