



Universidad Autónoma del Estado de México
Facultad de Economía



Licenciatura en Actuaría

Predicción de la tasa inflacionaria en México a través del Índice Nacional de Precios al Consumidor: 1969-2021

Tesis para obtener el título de Licenciados y Licenciada en Actuaría

Presentan:

Alvarez Navarro Jorge Arturo

Díaz Gaona Ricardo

Monroy Cruz María del Consuelo

Asesor:

Dr. En E.P. Daniel Lozano Keymolen

Revisores:

Dr. En E.P. Sergio Cuauhtémoc Gaxiola Robles Linares

Dra. En C.E.A. Natalie Ramírez Carmona

Marzo 2023

Índice

Introducción.....	1
Preguntas de Investigación	3
Hipótesis de la Investigación	3
Objetivo general	4
Objetivos específicos.....	4
Justificación de la investigación.....	4
Capítulo 1. Marco conceptual.....	6
1.1 El índice Nacional de Precios al Consumidor	6
1.1.1 Componentes del INPC.....	6
1.1.2 Cambios del año base del INPC.....	8
1.1.3 El cambio de base del 2018.....	8
1.2 El salario mínimo	9
1.2.1 El salario Mínimo en 2020	9
1.2.2 Componentes del salario	10
1.2.3 Determinación del salario mínimo en México	10
1.3 La Demanda	11
1.4 Tipo de cambio.....	13
Capítulo 2. La inflación en México, 1969 – 2022	15
2.1 Elementos definitorios de la Inflación en México.....	15
2.2 La Inflación en México: 1960-1994	19
2.3 La Inflación en México: 1994-2022	22
2.4 Factores asociados a la Inflación en México.....	32
Capítulo 3. Metodología de la investigación	34
3.1 Fuente de datos	34
3.2 Variables de investigación	35
3.3 Técnica de investigación.....	35
3.3.1 Modelo ARIMA (p,d,q).....	37
3.3.2 Modelo ARIMAX.....	39
3.3.3 Modelos ARCH y GARCH.....	40
3.3.4 Modelo de Redes neuronales	46
3.4 Comparación de modelos: SARIMA-GARCH; SARIMAX y Redes Neuronales	53

Capítulo 4. Resultados de la investigación	56
4.1 Modelo SARIMA.....	56
4.2 Modelo SARIMAX.....	86
4.3 Modelo de Red Neuronal.....	110
4.3.1. Perceptrón múltiple utilizando las expectativas y el INPC al tiempo t.....	111
4.3.2 Modelo con expectativas y dummies sobre los meses del año.....	115
4.3.3. Modelo utilizando únicamente el INPC al tiempo t.....	120
4.3.4. Modelo perceptrón ajustado en base a la experiencia de los modelos anteriores.....	124
4.3.5. Red Neuronal Autorregresiva.....	129
Discusión y Conclusiones	139
Limitaciones de la tesis.....	141
Futuras líneas de investigación.....	143
Anexos	144

Índice de tablas

Tabla 1.1. Diferencias en el INPC según el año base.	8
Tabla 1.2. Procedimiento para la determinación del salario mínimo	11
Tabla 2.1. Índice Nacional de Precios al Consumidor (Variación en por ciento)	25
Tabla 4.1. Medidas de error para las 4 ecuaciones candidatas	73
Tabla 4.2. Coeficientes y significancia del modelo $SARIMA(3, 1, 3)(1, 1, 2)$	1274
Tabla 4.3. Coeficientes y significancia del modelo $SARIMA(3, 1, 3)(0, 1, 2)$	1274
Tabla 4.4. Coeficientes y significancia de estos del modelo $SARIMA(3, 1, 3)(0, 1, 1)$	12 74
Tabla 4.5. Coeficientes y significancia de estos del modelo $SARIMA(1, 1, 2)(0, 1, 1)$	12 75
Tabla 4.6. Pronósticos para el año 2022 en escala logarítmica.	76
Tabla 4.7. Pronósticos para el año 2022 en su escala normal.	77
Tabla 4.8. Posibles valores de c según el grado de probabilidad	81

Tabla 4.9. Pronósticos del INPC en México, así como los intervalos de confianza construidos con (1) al 80 y 95%	83
Tabla 4.10. Pronósticos de la inflación mensual anualizada usando (2) en México, así como los intervalos de confianza construidos con (1) al 80 y 95%	85
Tabla 4.11. Comparación de las medidas de error MAE, RMSE y MAPE y otros criterios, entre las ecuaciones planteadas con diferentes variables exógenas construidas con el conjunto de entrenamiento y prueba	104
Tabla 4.12. Coeficientes y significancia del modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.	104
Tabla 4.13. Coeficientes y significancia del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.	105
Tabla 4.14. Coeficientes y significancia del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.	105
Tabla 4.15. Pronósticos del INPC para el año 2022 con el modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$	108
Tabla 4.16. Pronósticos para el año 2022 con el modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$	109
Tabla 4.17. Pronósticos de la inflación mensual anualizada en México con las variables exógenas salario mínimo, oferta y demanda agregada y expectativas inflacionarias.	109
Tabla 4.18. Predicción del perceptrón vs el INPC real	115
Tabla 4.19. Tabla de insumo que se utiliza para la estacionalidad de los meses	115
Tabla 4.20. Tabla de insumo que se utiliza para la estacionalidad de los meses	116
Tabla 4.21. Predicción del perceptrón vs el INPC real	119

Tabla 4.22. Predicción del perceptrón vs el INPC real	123
Tabla 4.23. Predicción del perceptrón vs el INPC real	127
Tabla 4.24. Predicción del perceptrón vs el INPC real utilizando la misma información de salida del perceptrón como entrada para el 2020.	127
Tabla 4.25. Predicción del perceptrón vs el INPC real utilizando la misma información de salida del perceptron como entrada para el 2022.	128
Tabla 4.26. Resumen de los modelos de Red Neuronal Ajustados al conjunto de datos.....	131
Tabla 4.27. Resumen de los modelos de Red Neuronal Ajustados al conjunto de datos.....	132
Tabla 4.28. Valores Pronosticados del INPC.....	133
Tabla 4.29. Valores Pronosticados del INPC.....	134
Tabla 4.30. Valores Pronosticados del INPC.....	135
Tabla 4.31. Medidas de error de las redes ajustadas	136
Tabla 4.32. Medidas de error de las redes ajustadas	136
Tabla 4.33. Pronósticos del INPC e Inflación mensual anualizada para 2022	137
Tabla A.1. Prueba de estacionariedad para el INPC en México.....	144
Tabla A.2. Prueba de estacionariedad para el INPC en México diferenciado completamente.....	144
Tabla A.3. Resumen estadístico para los coeficientes del modelo GARCH (1,1) ajustado a los residuales de la ecuación candidata.....	144
Tabla A.4. Prueba de estacionariedad del INPC	144
Tabla A.5. Prueba de estacionariedad del INPC conjunto de entrenamiento con transformación y primeras diferencias.	144
Tabla A.6. Prueba de estacionariedad de las variables exógenas	145

Índice de gráficas

Gráfico 1.1. Comportamiento de la inflación en relación con las variables...	13
Gráfico 2.1. Comportamiento en el nivel de precios condicionado a un aumento en la producción.....	16
Gráfico 2.2. Evolución de la inflación en México de febrero de 1969 a diciembre 2021.	17
Gráfico 2.3. Inflación mensual al cuadrado: febrero 1969 a diciembre 2021.	18
Gráfico 2.4. Inflación anual en México de enero 1988 a diciembre 1995.....	23
Gráfico 2.5. INPC vs INPC subyacente (Variación anual en por ciento).	28
Gráfico 3.1. Representación gráfica de una red neuronal alimentada hacia adelante.	47
Gráfico 3.2. Función de activación RELU.....	49
Gráfico 3.3. Función de activación Sigmoide.	49
Gráfico 3.4. Función de activación Tangente hiperbólica.	50
Gráfico 3.5. Desarrollo de la metodología Box & Jenkins para la elección de un modelo de serie de tiempo.	54
Gráfico 4.1. Índice Nacional de Precios al Consumidor: enero de 1969 a diciembre 2021.	56
Gráfico 4.2. Valores del INPC vs sus primeros 16 rezagos.	59
Gráfico 4.3. ACF del INPC.....	59
Gráfico 4.4. Transformación Logarítmica del INPC en México (Conjunto de entrenamiento).	61
Gráfico 4.5. Primeras diferencias del conjunto de entrenamiento.....	61
Gráfico 4.6. ACF de las primeras diferencias del conjunto de entrenamiento con transformación logarítmica.	63
Gráfico 4.7. PACF de las primeras diferencias del conjunto de entrenamiento con transformación logarítmica.	63

Gráfico 4.8. Descomposición del conjunto de entrenamiento desde el enfoque Multiplicativo.	64
Gráfico 4.9. Diferencias estacionales aplicadas al conjunto de entrenamiento.	65
Gráfico 4.10. Proceso diferenciado completamente aplicado al conjunto de entrenamiento.....	65
Gráfico 4.11. ACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento.....	66
Gráfico 4.12. PACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento.....	66
Gráfico 4.13. Residuales del modelo $SARIMA(3, 1, 3)(1, 1, 2)_{12}$.....	69
Gráfico 4.14. Residuales del modelo $SARIMA(3, 1, 3)(0, 1, 2)_{12}$.....	69
Gráfico 4.15. Residuales del modelo $SARIMA(3, 1, 3)(0, 1, 1)_{12}$.....	70
Gráfico 4.16. Residuales del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$.....	70
Gráfico 4.17. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(1, 1, 2)_{12}$.	71
Gráfico 4.18. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(0, 1, 2)_{12}$.	72
Gráfico 4.19. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(0, 1, 1)_{12}$.	72
Gráfico 4.20. Comparativa: Conjunto de prueba vs modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$.	73
Gráfico 4.21. Residuales del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$ tomando todo el conjunto de datos disponible.....	76
Gráfico 4.22. Residuales al cuadrado del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$ tomando todo el conjunto de datos disponible.....	78
Gráfico 4.23. ACF de los Residuales al cuadrado del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$ tomando todo el conjunto de datos disponible.....	79

Gráfico 4.24. Ajuste de la ecuación GARCH (1,1) a los residuales de la ecuación candidata.....	80
Gráfico 4.25. Intervalos de confianza al 80% para el pronóstico del INPC en México.	82
Gráfico 4.26. Intervalos de confianza al 95% para el pronóstico del INPC en México.	82
Gráfico 4.27. Intervalos de confianza al 80% para el pronóstico de la inflación en México usando (2) y (1).....	85
Gráfico 4.28. Intervalos de confianza al 95% para el pronóstico de la inflación en México usando (2) y (1).....	86
Gráfico 4.29. Índice Nacional de Precios al Consumidor: enero de 2000 a diciembre de 2021.	87
Gráfico 4.30. ACF del INPC 2000-2021.....	88
Gráfico 4.31. INPC conjunto de entrenamiento con transformación logarítmica.	89
Gráfico 4.32. INPC conjunto de entrenamiento transformado con primeras diferencias.	89
Gráfico 4.33. ACF del conjunto de entrenamiento transformado y con primeras diferencias.	90
Gráfico 4.34. PACF del conjunto de entrenamiento transformado y con primeras diferencias.	90
Gráfico 4.35 Descomposición del conjunto de entrenamiento transformado.....	91
Gráfico 4.36. INPC conjunto de entrenamiento transformado añadiendo diferencias estacionales.....	91
Gráfico 4.37. INPC conjunto de entrenamiento transformado añadiendo diferencias estacionales y primeras diferencias.....	92
Gráfico 4.38. ACF del conjunto de prueba transformado y diferenciado completamente.....	93

Gráfico 4.39. PACF del conjunto de entrenamiento transformado y diferenciado completamente.....	93
Gráfico 4.40. Comportamiento del conjunto de entrenamiento transformado de la variable Tipo de cambio.....	94
Gráfico 4.41. Primeras diferencias y diferencias estacionales al conjunto transformado de entrenamiento de la variable Tipo de cambio.....	94
Gráfico 4.42. Comportamiento del conjunto de entrenamiento transformado de la variable Salario mínimo.	95
Gráfico 4.43. Primeras diferencias y diferencias estacionales al conjunto de entrenamiento transformado de la variable Salario mínimo.	95
Gráfico 4.44. Comportamiento del conjunto de entrenamiento transformado de la variable Oferta y Demanda agregadas.	96
Gráfico 4.45. Conjunto de prueba transformado de la Variable Oferta y Demanda agregadas con primeras diferencias y diferencias estacionales.....	97
Gráfico 4.46. Residuales del modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	100
Gráfico 4.47. Residuales del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	100
Gráfico 4.48. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	101
Gráfico 4.49. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con variables salario mínimo y oferta y demanda agregada.	101
Gráfico 4.50. Comparación del conjunto de prueba vs modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	102
Gráfico 4.51. Comparación del conjunto de prueba vs modelo $SARIMA(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	102

Gráfico 4.52. Comparación del conjunto de prueba vs modelo $SARIMA(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.....	103
Gráfico 4.53. Comparación del conjunto de prueba vs modelo $SARIMA(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo y oferta y demanda agregada.	103
Gráfico 4.54. Residuales del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias tomando todo el conjunto de datos disponible.....	107
Gráfico 4.55. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias tomando todo el conjunto de datos disponible.....	107
Gráfico 4.56. Perceptrón multicapa conformado por dos neuronas de entrada, dos capas ocultas de 4 y 3 neuronas con una neurona en la capa de salida.	111
Gráfico 4.57. Grafica de magnitud de pérdida en el entrenamiento.....	113
Gráfico 4.58. Grafica de comparación entre datos originales, entrenamiento y prueba.	114
Gráfico 4.59. Comparativa: INPC y datos simulados por el perceptrón.	114
Gráfico 4.60. Perceptrón (14,6,2,1).....	117
Gráfico 4.61. Magnitud de perdida en el entrenamiento.	117
Gráfico 4.62. Comparación entre datos originales, entrenamiento y prueba.	118
Gráfico 4.63. Comparativa: INPC y datos simulados por el perceptrón.	119
Gráfico 4.64. Perceptrón (1,10,5,2,1).....	120
Gráfico 4.65. Magnitud de pérdida en el entrenamiento.	121
Gráfico 4.66. Comparación entre datos originales, entrenamiento y prueba.	122
Gráfico 4.67. Comparativa: INPC y datos simulados por el perceptrón.	122
Gráfico 4.68. Perceptrón (1,10,7,2,1).....	124

Gráfico 4.69. Magnitud de pérdida en el entrenamiento	125
Gráfico 4.70. Comparación entre datos originales, entrenamiento y prueba.	126
Gráfico 4.71. Comparativa: INPC y datos simulados por el perceptrón.	126
Gráfico 4.72. ACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento sin transformación logarítmica.....	130
Gráfico 4.73. PACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento sin transformación logarítmica.....	131
Gráfico 4.74. Inflación mensual anualizada pronosticada con el modelo $NNAR(11, 2, 7)$.....	138

Introducción

De acuerdo con la definición propuesta por el Banco de México (BANXICO), la inflación es un fenómeno observado dentro de la economía de un país. Dicha variable tiene relación con el aumento generalizado y a la vez sostenido en el tiempo de los precios en la mayor parte de los bienes y servicios que se comercian dentro de sus mercados: su estudio y entendimiento resulta ser vital al momento de conocer el estado actual de una economía (Banxico, Banxico-Educa, s.f.).

La inflación se puede explicar en relación con su comportamiento a través del tiempo y también a través de otras variables, que dentro de una economía su movimiento en el tiempo pueda desembocar en una inflación alta o en la regulación de esta. Las variables que serán consideradas en esta investigación son el tipo de cambio (paridad peso-dólar), el salario mínimo, la oferta y demanda agregada y las expectativas inflacionarias recabadas de la encuesta realizada a expertos por el Banco de México (BANXICO, Salario mínimo e inflación, 2016).

El Banco de México realizó un estudio demostrando que el incremento del salario mínimo incrementa de manera directa a la inflación y al nivel de precios durante los primeros 12 meses, y se relaciona al tipo de cambio en el largo plazo (BANXICO, Salario mínimo e inflación, 2016).

Un aumento en la demanda por parte de la población lleva al aumento de la oferta y de los precios de producción, que si se generaliza en los diferentes sectores da como resultado un alza en la inflación (Banxico, 2020).

De acuerdo con las expectativas de la inflación los agentes económicos toman decisiones dependiendo la tasa que se espere en el futuro, se realizan ajustes a las políticas monetarias para cumplir con el objetivo establecido por el Banco de México (BANXICO, Evolución de las Expectativas de Inflación de Largo Plazo, 2019).

En pleno 2022 puede observarse que el incremento de los precios a nivel general en México ha sido sostenido, por lo que es posible afirmar que actualmente existe un periodo inflacionario. Las causas de este fenómeno resultan ser diversas, sin embargo, las más sobresalientes tienen relación con la pandemia por coronavirus (COVID-19) la cual ocasionó un crecimiento acelerado de la inflación debido a que importantes cadenas de

valor se rompieron, ocasionando una disminución importante en el flujo de comercios dentro del país (Instituto Belisario Domínguez, 2020, p.1-2).

Por otra parte, el conflicto bélico entre Rusia y Ucrania propició el aumento de los precios debido a las sanciones impuestas a dichos países a nivel de sus exportaciones e importaciones alterando en buena medida el comportamiento dentro de los mercados financieros (Informe trimestral, 2022).

Debido al problema actual con el comportamiento de la inflación, surge la necesidad de tener una herramienta que permita tener un amplio panorama del comportamiento futuro inflacionario en México. Entonces, esta tesis tiene un enfoque metodológico en el cual se muestra el ajuste de diversos modelos estadísticos que permiten pronosticar el comportamiento futuro de la inflación a través del Índice Nacional de Precios al Consumidor (INPC). Estos modelos encuentran respaldo en la literatura, pero se agrega un nuevo enfoque de análisis mediante la aplicación de una técnica de Redes Neuronales.

Dentro de esta investigación se desarrollan y comparan 3 diferentes metodologías de pronóstico las cuales son: modelos Autorregresivos Integrados y de Medias Móviles Estacionales (SARIMA en inglés), el Autorregresivo Integrado y de Medias Móviles Estacional con Variables Explicativas (SARIMAX) que debe su nombre a la implementación de variables exógenas que expliquen los niveles de inflación en nuestro país y, el modelo de Red Neuronal que aprende el comportamiento de la variable de interés.

La implementación de 3 metodologías distintas tiene la finalidad de conocer aquel modelo que pronostica mejor en un corto plazo (1 año), para lo cual se compararon sus medidas de precisión como el Error Cuadrático Medio (RMSE en inglés), el Error Absoluto Medio (MAE en inglés) y el Error Absoluto Porcentual medio (MAPE en inglés), así como propiedades estadísticas importantes dentro de un modelo de serie de tiempo para aquellos modelos considerados como tal (SARIMA-GARCH, SARIMAX). Para el tema de la Red Neuronal se compararon las medidas de precisión y el apego a la realidad de los pronósticos obtenidos que es un indicador con cierto grado de subjetividad.

Es importante decir que los modelos desarrollados cubren distintos periodos de tiempo dadas las diferencias en la disponibilidad de datos. En el caso de las aplicaciones de los modelos SARIMA-GARCH y de Redes Neuronales se emplearon datos de 1969 a 2021, así mismo utilizar el periodo comprendido de enero de 2000 a diciembre de 2021 para el modelo SARIMAX y determinar qué modelo se adapta mejor a la realidad de este indicador.

El objetivo de realizar lo anterior, es mostrar las capacidades técnicas que cada modelo presenta y que puedan emplearse como referentes para el desarrollo de pronósticos. Así, con el proyecto se busca determinar el mejor modelo en términos estadísticos y conciliado conceptualmente que ayudará a tomar decisiones en un corto o mediano plazo y de esta forma poder predecir con una mejor exactitud la tasa inflacionaria. Entonces, se exponen las siguientes preguntas de investigación que son la guía de este proyecto.

Preguntas de Investigación

- 1) ¿Es posible proyectar el valor de la inflación en México según los efectos de variables como el salario mínimo, demanda, tipo de cambio y las expectativas inflacionarias?
- 2) ¿Puede definirse un modelo de pronósticos de la inflación que resulten en valores próximos a los proyectados por el Banco de México?

Hipótesis de la Investigación

- 1) Puede pronosticarse el INPC para después aproximarse a la inflación de forma eficaz con base en el comportamiento de las variables salario mínimo, demanda, tipo de cambio y las expectativas inflacionarias.
- 2) Al comparar las proyecciones de la inflación en México mediante modelos SARIMA-GARCH, SARIMAX o de redes neuronales, se encontrará el modelo que mejor se adecue a los valores reales publicados por el Banco de México.

Objetivos de investigación

Objetivo general

Pronosticar la tasa de inflación a través de su indicador económico el Índice Nacional de Precios al Consumidor en el periodo 1969 a 2021 para los modelos SARIMA-GARCH y Redes Neuronales, así mismo utilizar el periodo comprendido de enero de 2000 a diciembre de 2021 para el modelo SARIMAX y determinar qué modelo se adapta mejor a la realidad de este indicador.

Objetivos específicos

- 1) Desarrollar teóricamente los conceptos esenciales de aquellas variables exógenas que intervienen en el modelo SARIMAX, así mismo explicar con cierto nivel de detalle la variable dependiente a pronosticar.
- 2) Analizar el comportamiento de la variable inflación a lo largo de la historia de nuestro país.
- 3) Explicar detalladamente las diferentes metodologías de pronóstico que serán utilizadas para proyectar la inflación mensual anualizada en México.
- 4) Desarrollar un modelo SARIMA que tome en cuenta el patrón estacional para realizar pronósticos del INPC y así mismo construir intervalos de confianza con un modelo GARCH que trate de manera adecuada la volatilidad representada en la serie, así mismo adecuar un modelo SARIMAX que tome en cuenta el patrón estacional de la serie y utilice las variables exógenas descritas en el marco conceptual. Por último, construir un modelo de redes neuronales para proyectar valores futuros del INPC para así aproximar los niveles de inflación en México.

Justificación de la investigación

Este trabajo de tesis tiene la función principal de predecir la tasa inflacionaria a través del INPC, utilizando diferentes modelos para poder estimar la mejor predicción con el menor número de errores posibles. Las fluctuaciones de la inflación pueden desencadenar problemas económicos graves como los vistos a lo largo de la historia del México contemporáneo (Véase el capítulo 2).

Para llevar a cabo el desarrollo y estudio comparativo de los modelos antes mencionados el presente trabajo se ha estructurado en 4 capítulos. En el primero de estos se explican a detalle las variables que mantienen una relación causal con la inflación, mismas que se han de utilizar al momento de llevar a cabo el ajuste para el modelo SARIMAX. En el segundo capítulo se explica la variable de interés (inflación) y pone en contexto al lector sobre las crisis inflacionarias que México ha vivido a lo largo de su historia. Para el capítulo 3 se explica la metodología de investigación utilizada para los pronósticos, en este se expone el fundamento teórico que respalda a las series temporales (SARIMA-GARCH, SARIMAX), así como las nuevas metodologías de pronóstico (Red Neuronal). En el capítulo cuarto se muestran los resultados, conclusiones y limitaciones de esta investigación confirmando que las nuevas metodologías de pronóstico (Red Neuronal Autorregresiva) proporciona una mayor precisión al momento de pronosticar la variable de interés (INPC e inflación).

De igual forma, en esta investigación se analizan y contrastan tres diferentes técnicas metodológicas para la realización de pronósticos de la variable inflación ya que otros trabajos revisados como Soriano (2020), Rodríguez (2020) y Munguía y Estrada (2018) han aplicado series de tiempo únicamente de tipo ARIMA estacional y modelo ARIMA GARCH.

Debido a nuestra formación académica como Actuarios somos capaces de realizar este tipo de análisis que permitan mejorar o ampliar la perspectiva en la creación de pronósticos sobre indicadores económicos importantes y relevantes, además, este trabajo sirve de ejemplo como referencia para otros alumnos que están estudiando actualmente.

Capítulo 1. Marco conceptual

El enfoque conceptual de esta tesis se basa en las concepciones económicas del Índice Nacional de Precios al Consumidor, el Salario, la Demanda, el Tipo de cambio y las Expectativas de inflación, conceptos que resultan ser clave a la hora de detallar las variables que intervienen en los modelos estadísticos y de series temporales usados en este trabajo. A continuación, de forma abreviada se exponen estos.

1.1 El índice Nacional de Precios al Consumidor

El Banco de México en su informe trimestral del 2018 define al Índice Nacional de Precios al Consumidor (INPC) como: *“un indicador diseñado específicamente para medir el cambio promedio de los precios en el tiempo, mediante una canasta ponderada de bienes y servicios representativa del consumo de las familias de México”* (BANXICO, Principales Elementos del Cambio base del INPC, 2018, págs. 55-56). El INPC como indicador mide el fenómeno económico que se conoce como inflación (BANXICO, 2002).

Para su construcción, a la canasta de bienes y servicios se le asignan distintas ponderaciones según la importancia de cada componente en el consumo, estos deben ser revisados con frecuencia para asegurar la representación que tengan en los hogares ya que tiende a cambiar, como dice el Banco de México *“en respuesta a varios factores, como los movimientos en precios relativos, la introducción de nuevos productos, la aparición de nuevos puntos de venta y los cambios en el ingreso, entre otros.”* (BANXICO, Principales Elementos del Cambio base del INPC, 2018).

1.1.1 Componentes del INPC

Los principales componentes del INPC se agrupan en ocho categorías, de acuerdo con la forma en que los consumidores distribuyen su gasto (BANXICO, 2002): Alimentos, bebidas y tabaco; Ropa, calzado y accesorios; Vivienda; Muebles, aparatos y accesorios domésticos; Salud y cuidado personal; Transporte; Educación y esparcimiento; Otros servicios.

El Instituto Nacional de Estadística y Geografía de México es quien construye este índice desde julio del 2011, actualmente reporta el orden siguiente de los componentes del INPC: (I NEGI, 2021)

- 1) **Subyacente.** Mercancías como alimentos, bebidas, tabaco y mercancías no alimentarias. Servicios como vivienda, educación y otros servicios.
- 2) **No subyacente.** Que incluye productos agropecuarios como las frutas, verduras y pecuarios. Energéticos y tarifas como son los propios energéticos y las tarifas autorizadas por el Gobierno.

Para calcular el INPC los precios de los productos considerados se agrupan en grupos casi homogéneos llamados genéricos, y tienen la menor ponderación del INPC. Cada mes se hace la recopilación de 170,000 productos específicos¹ de que se agrupan en 313 genéricos (BANXICO, 2002).

La investigación de los productos que se utilizan en el cálculo se realiza en tiendas, comercios y de prestadores de servicios que son preferidos por los consumidores de las 46 ciudades de cotización, se seleccionan ciertos productos cuyo precio² se les hace un seguimiento recurrente (BANXICO, 2002).

El cálculo del INPC considera cuatro estratos diferentes de acuerdo con el nivel de ingreso de los hogares según los salarios mínimos disponibles en el mismo. Estos grupos son:

- a) Estrato 1: Hogares con ingresos menores que un salario mínimo.

¹ El sistema del INPC se integra de 46 ciudades y áreas metropolitanas agrupadas en siete regiones. A su vez, por su tamaño, las ciudades se clasifican en pequeñas, medianas y grandes. De este modo se calculan índices de precios para cada una de las siete regiones en que se divide el territorio nacional y para las 46 ciudades que conforman el sistema, así como para cada tamaño de localidad. Cabe señalar que al menos en una ciudad por estado se recoge información para el cálculo del INPC. De esta manera se asegura la representatividad espacial del INPC (BANXICO, 2002).

² La selección de los servicios se realiza con base en la opinión de los informantes que escogen a los más representativos. Los precios que se recaban pasan por un proceso de revisión y, si es necesario, de verificación. Así, se garantiza que las cotizaciones que intervienen en el cálculo del INPC son los precios vigentes en el mercado (BANXICO, 2002).

- b) Estrato 2: Hogares con ingresos entre uno y tres salarios mínimos.
- c) Estrato 3: Hogares con ingresos entre tres y seis salarios mínimos.
- d) Estrato 4: Hogares con ingresos mayores que seis salarios mínimos.

Para cada uno de los grupos anteriores se calcula un índice, que considera las preferencias de consumo en cada estrato (BANXICO, 2002).

1.1.2 Cambios del año base del INPC

El INPC ha estado sujeto a seis cambios de año base desde 1969, y se han modificado tanto las ciudades consideradas, como el número de productos y servicios cotizados.

En la Tabla 1.1 se exponen las diferencias en el INPC según el año base en el periodo 1968 y la quincena 2 de julio de 2018.

Tabla 1.1. Diferencias en el INPC según el año base.

Periodo base	Ciudades	Número de productos y servicios cotizados
1968	7	5,100
1978	16	7,100
1980	35	9,700
1994	46	43,400
2da quincena de junio 2002	46	58,200
2da quincena de diciembre 2010	46	83,500
2da quincena de julio 2018	55	120,454

Fuente: Elaboración propia con el informe trimestral de 2018 de BANXICO (BANXICO, Principales Elementos del Cambio base del INPC, 2018).

1.1.3 El cambio de base del 2018

Los bienes y servicios que se consideraron para la nueva canasta se eligieron según la Encuesta Nacional de Gastos de los Hogares (ENGASTO) del 2012 y 2013 y la Encuesta

Nacional de Ingresos y Gastos de los Hogares (ENIGH) del 2014 (BANXICO, Principales Elementos del Cambio base del INPC, 2018).

De la nueva canasta 265 de los genéricos se mantienen igual, se incorporan 10 nuevos genéricos, 2 desaparecen, 9 se desagregan en 19 y 7 se fusionan en 5, teniendo un total de 299 genéricos a comparación de los 283 anteriores (BANXICO, Principales Elementos del Cambio base del INPC, 2018).

1.2 El salario mínimo

La Comisión Nacional de los Salarios Mínimos (CONASAMI) define al salario mínimo como la cantidad menor que debe recibir en efectivo el trabajador por los servicios prestados en una jornada de trabajo (Rambiola, 2019). En forma institucional, la CONASAMI en su reporte de julio del 2011 define al salario mínimo:

“Con la promulgación de la Constitución General de la República publicada en el Diario Oficial de la Federación del 5 de febrero de 1917, específicamente en el artículo 123, fracción VI bajo el principio de que el salario mínimo deberá ser suficiente para satisfacer las necesidades normales de la vida del obrero, su educación y sus placeres honestos, considerándolo como jefe de familia...”. Se determina, asimismo, en la fracción VIII, que el salario mínimo no podrá ser objeto de embargo, compensación o descuento alguno” (CONASAMI, 2011, pág. 1)

En México el salario mínimo tiende a incrementar una vez al año, debido a la inflación que existe en el país misma que va reduciendo el poder adquisitivo de los sueldos y por ello se necesita un porcentaje de mejora. Normalmente el aumento lo da la CONASAMI entre el 1 y el 31 de diciembre, pero entra en vigor hasta enero del siguiente año (Rambiola, 2019).

1.2.1 El salario Mínimo en 2020

El Consejo de Representantes de la CONASAMI acordó para finales de 2019 aumentar el salario mínimo pasando de 102.68 a 123.22 pesos diarios. El nuevo salario mínimo se

determinó partiendo del salario anterior de 2019, al cual se le adicionaron 14.67 pesos a través del llamado Monto Independiente de Recuperación (MIR)³ (Rambiola, 2019).

1.2.2 Componentes del salario

Según (Rambiola, 2019) los componentes del salario son:

- a) El aguinaldo: se considera al pago anual de una porción del salario, son mínimo 15 días de salario.
- b) Prima dominical: Para los que laboran el domingo y tienen otro día de descanso en la semana, es un 25% más a su salario.
- c) Horas extras: Si se exceden las 8 horas diarias, nueve horas deben pagarse con un 100% adicional de sueldo y después de las nueve horas semanales se recibe un 200% de salario por cada hora extra correspondiente a lo que se le paga la misma.
- d) Vacaciones y prima vacacional: Días obligatorios de descanso que se cobran como días de trabajo laboral, adicionando una vez al año un 25% al salario prima vacacional.
- e) Propinas: Para empleados de hoteles, restaurantes, bares, entre otros.
- f) Participación de utilidades: Por Ley⁴, cada trabajador tiene derecho a recibir una parte de las utilidades de la empresa de acuerdo con el número de días que laboró y al monto de su salario.

1.2.3 Determinación del salario mínimo en México

El artículo 123 de la Constitución General de la República informa que el salario mínimo deberá ser suficiente para satisfacer las necesidades normales de un jefe de familia en el orden material, social y cultural, así como también para proveer la educación obligatoria de sus hijos si es que los tuviese. Anualmente se le adiciona un porcentaje que no supera

³ El poder judicial de Chiapas en su compilación de Legislación y Jurisprudencia dice que el Monto Independiente de Recuperación “es la aplicación o incremento al salario por concepto de MIR” y es para trabajadores asalariados cuyo ingreso diario sea el mínimo general, entró en vigor en 2017.

⁴ Artículo 122 de la Ley Federal del Trabajo

al INPC, dicho porcentaje tiene una relación directa con el tema de la inflación y el poder adquisitivo de las personas (Rambiola, 2019).

En la Tabla 1.2 se resume el proceso que lleva a cabo la CONASAMI para la fijación del nuevo salario mínimo.

Tabla 1.2. Procedimiento para la determinación del salario mínimo

Primero	De acuerdo con la fracción III del artículo 561 y en el artículo 562 de la Ley Federal del Trabajo la Dirección Técnica de la CONASAMI tiene el deber de investigar las condiciones generales de la economía del país, las variaciones en el costo de la vida de las familias, y el impacto de los salarios definidos para 2020 sobre el empleo y la estructura salarial.
Segundo	La Dirección técnica obtiene información de Instituciones oficiales sobre la evolución de las condiciones del mercado laboral, las estructuras salariales, y recibe dentro del plazo legal los informes y sugerencias hechos por los representantes de patrones y trabajadores.
Tercero	El Consejo de Representantes recibe el informe de la Dirección Técnica, así como los informes mensuales que presentó al Consejo que fueron aprobados por el mismo.
Cuarto	Se establecen los salarios mínimos por área geográfica (zona libre de la frontera norte y el resto del país) y por profesiones, oficios y trabajos especiales.
Quinto	Los salarios se aprueban por medio del voto aprobatorio de los Representantes Propietarios y Suplentes de los trabajadores y del presidente de la CONASAMI participantes en la resolución. La resolución de los salarios mínimos se publicará en el Diario Oficial de la Federación a más tardar el 31 de diciembre de cada año.

Elaboración propia con la información del Diario Oficial de la Federación (2020).

1.3 La Demanda

Una de las variables de estudio que tienen una gran relación teórica con la tasa inflacionaria en México es la demanda, aunque cabe mencionar que todo está relacionado ya que se crea todo un ciclo entre la demanda, precios, salarios y tasas inflacionarias. Se sabe que la inflación obedece las leyes de oferta y demanda, lo que quiere decir que “si la

demanda de bienes excede la capacidad de producción o importación de bienes, los precios tienden a aumentar”⁵.

Un dato importante para analizar en este trabajo es que el sector real de la economía es más vulnerable a cambios en la oferta temporales por la razón de que al tener un aumento temporal en algunos insumos la forma de contrarrestar esto es aumentar la tasa de interés para así evitar el aumento de precios en los bienes finales contrayendo la demanda agregada (Garcia, 2002). Según García 2003,

“Cuando las presiones de inflación se originan por el lado de la demanda, derivadas de un exceso de gasto, la respuesta de la política monetaria es hacia un aumento en la tasa de interés tanto por el efecto de la brecha de inflación como por el de la brecha del producto” (Torres García, 2003, pág. 12).

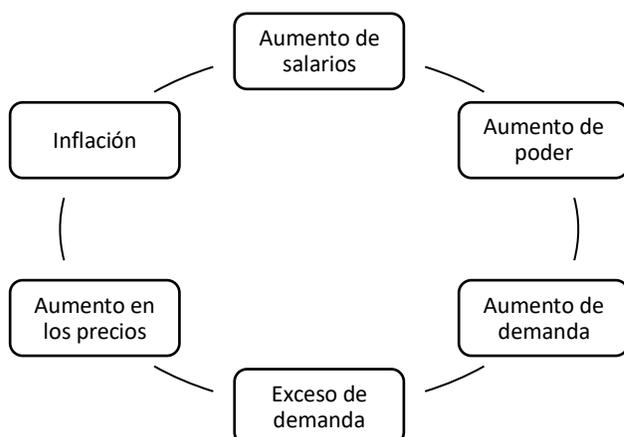
La demanda agregada se considera como un determinante de corto plazo en la inflación, ya que cuando se tiene un exceso de demanda en un mercado donde hay pocos productos las empresas se ven obligadas a incrementar su precio lo que causa una inflación (Banxico, 2020). Así, la inflación va muy ligada a la parte de la oferta y demanda ya que se tienen que hacer los cálculos necesarios para encontrar el precio de equilibrio de los bienes en cuestión, pero para poder entender esto es fundamental poder explicar cómo funciona la oferta y la demanda en los precios.

Para explicar el principio de la Ley de Oferta y Demanda sería preciso definir la idea principal en la que se basa, principalmente se debe encontrar un punto de equilibrio entre lo que es la oferta y la demanda de los bienes, ya que cuando hay un bien en el mercado a un precio en específico, los consumidores van a buscar adquirirlo de acuerdo con sus posibilidades buscando maximizar su utilidad (Blanchard, Amighini , & Giavazzi, 2012).

Una vez explicado lo anterior se puede llegar al punto principal del círculo constante que hay entre la demanda de los productos, los salarios y la inflación mostrado en el siguiente gráfico.

⁵ Véase la sección 2.1.

Gráfico 1.1. Comportamiento de la inflación en relación con las variables



Fuente: Elaboración propia con la información de macroeconomía teoría y políticas (2007).

1.4 Tipo de cambio

El tipo de cambio puede definirse, según Banxico, como aquella referencia que es utilizada dentro de los mercados para saber con cuántas monedas nacionales podemos adquirir una moneda extranjera, así mismo puede definirse como la cantidad de monedas nacionales que se reciben a cambio de vender una moneda extranjera. Este tema por sí solo puede decirse que es demasiado extenso pues todo depende principalmente de los regímenes cambiarios de cada país, en este caso Banxico define los siguientes:

- 1) Régimen cambiario fijo: En este sistema el tipo de cambio se determina o queda “fijo” por las autoridades monetarias que intervienen en el mercado para comercializar con las divisas, es decir, compran o venden divisas para asegurar la estabilidad del tipo de cambio.
- 2) Régimen flexible o flotante: El precio de la divisa es determinada según las condiciones de oferta y demanda de los países, aquí los precios por cada moneda no son intervenidos por las autoridades monetarias.
- 3) Régimen de bandas cambiarias: En este sistema se combinan los dos regímenes anteriores con la finalidad de mantener estable el precio de la moneda en cuestión, en términos sencillos las autoridades monetarias tienden a fijar un intervalo en donde el precio del tipo de cambio puede variar libremente, si dicho intervalo de precios es violado, las autoridades monetarias intervienen para regular los precios que han sido modificados por la oferta o demanda, según sea el caso.

Existen diversos casos que pueden afectar positiva o negativamente una economía, estos casos pueden ser cuando la moneda en cuestión se aprecia o se deprecia, para el caso en que la moneda se aprecia estamos frente a una situación en la que la moneda extranjera se hace más barata frente a la nacional, por otro lado, se dice que la moneda se deprecia cuando la moneda extranjera se aprecia y aumenta su valor, esto quiere decir que ahora cuesta más comprarla a comparación de antes.

Dichos cambios son importantes pues pueden afectar directamente al tema inflacionario; estos cambios, por ejemplo, se ven reflejados en el tema de la adquisición de bienes y servicios si se supone una economía en donde los productos que se venden dentro de ella se adquieren directamente en el extranjero y posteriormente se comercializan en dicha economía al subir el tipo de cambio a favor del extranjero los precios de los bienes y servicios adquiridos para la economía que los compra subirán y por consiguiente será más caro adquirirlos. Por el contrario, donde la economía en cuestión es la que vende los productos y/o servicios al extranjero, quizá ahora para la economía extranjera será un poco difícil el adquirir dichos productos, y esto repercutiría directamente en el crecimiento de las exportaciones de la economía en cuestión, haciendo que el crecimiento sea lento y desproporcionado a comparación de cuando el tipo de cambio era regular (Monex, 2020).

Para respaldar el punto anterior algunos trabajos han mostrado la relación que tiene el tipo de cambio con la inflación. De esta manera, se estima que los cambios en el tipo de cambio o mejor dicho cuando hay una depreciación del tipo de cambio es cuando se ejercen presiones inflacionarias (López, 2019).

Expuesto el marco conceptual del proyecto, en el siguiente capítulo se describe el objeto de investigación de la tesis.

Capítulo 2. La inflación en México, 1969 – 2022

El objetivo de este capítulo es definir y describir el comportamiento de la inflación en México durante el periodo de análisis de esta tesis.

2.1 Elementos definitorios de la Inflación en México

La inflación es un fenómeno que afecta a toda la economía y es una variable cuya importancia radica en la toma de decisiones. Esta variable puede definirse como el aumento generalizado y a la vez sostenido en los precios dentro de una economía (Heath J. , 2012). El estudio de la inflación es de vital importancia para poder entender de mejor manera el estado actual de la misma (Heath J. , 2012). Autores como Blanchard (2017) coinciden con la definición propuesta por Heath para el término inflación, que en resumidas cuentas se refiere al incremento de manera general de los precios. Es posible hacer una distinción en el concepto de esta variable, pues los precios no crecen por sí mismos, sino que estos incrementan su valor debido a una tasa, dicha tasa es conocida como tasa de inflación y se define como aquella medida a la que sube el nivel de precios.

Es importante señalar que la definición de inflación tiene dos condiciones que son de suma importancia:

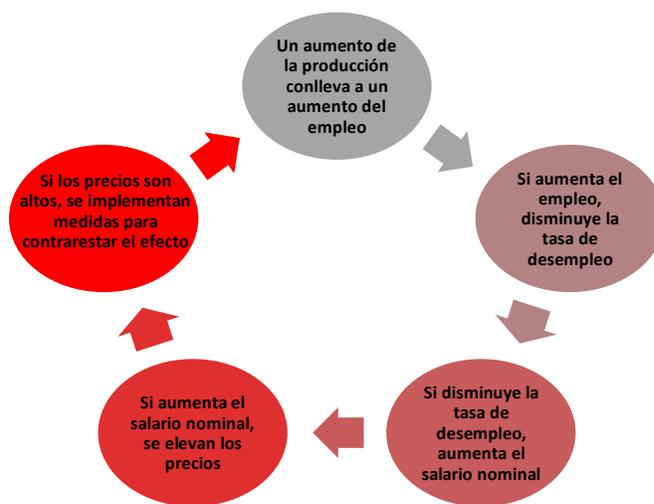
1. Debe haber un aumento de manera general en los precios en una economía y no solo un aumento en determinados productos.
2. El incremento en los precios debe ser sostenido, esto quiere decir que los precios incrementarán a lo largo de un periodo en específico.

En términos aclaratorios, es importante no confundir a la inflación con los cambios relativos en los precios dentro de una economía, ya que dichos cambios expresan el comportamiento en el precio de diferentes productos respecto a otros. Acorde a Heath (2012), sí existe una economía en donde la misma se regula por la oferta y la demanda, los precios tenderán a ser variables, es decir, éstos tienden a tener un comportamiento de manera libre y debido a eso alertan sobre la escasez o la abundancia para un producto en particular.

Cuando se habla del fenómeno de aumento generalizado y sostenido dentro de los precios, generalmente tiene que ver con las variaciones de la oferta y demanda agregada, las cuales pueden definirse como sigue:

1.- Oferta agregada (OA): Cantidad total de determinado bien que circula en una economía, generalmente son aquellos productos ofertados por las empresas, es importante señalar que esta definición sostiene una relación importante con el nivel de precios, pues la relación de oferta agregada nos muestra como los precios se ven afectados si es que se modifica el nivel de producción (Blanchard, Amighini , & Giavazzi, 2012). En este sentido, si el nivel de producción aumenta, se espera exista un aumento en el nivel de precios, este proceso se resume en el Gráfico 2.1.

Gráfico 2.1. Comportamiento en el nivel de precios condicionado a un aumento en la producción.

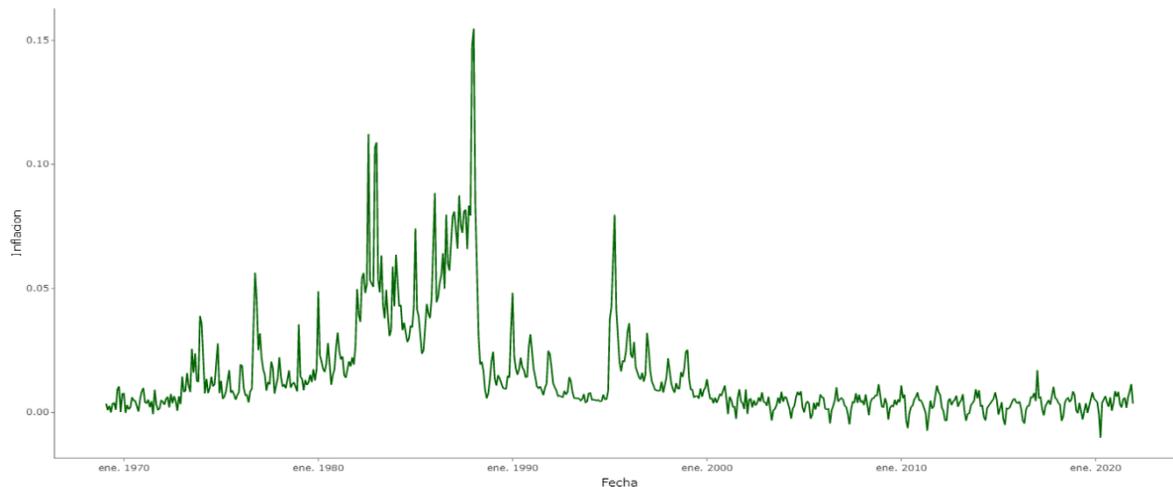


Fuente: Elaboración propia a partir de la información encontrada en (Blanchard, Amighini , & Giavazzi, 2012).

2.- Demanda agregada (DA): Se puede definir como el total de bienes y/o servicios que circulan en una economía para su consumo, la relación de demanda agregada nos muestra el efecto que tiene el nivel de precios en la producción a partir de las condiciones de equilibrio de mercado (Blanchard, Amighini , & Giavazzi, 2012).

En México, el efecto de la inflación ha jugado un papel importante en ciertos periodos de tiempo a lo largo de su historia, debido a lo anterior, en el Gráfico 2.2 se presenta la evolución de la inflación mensual en México de febrero de 1969 a diciembre de 2021.

Gráfico 2.2. Evolución de la inflación en México de febrero de 1969 a diciembre 2021.



Fuente: Elaboración propia utilizando ggplotly() en R.

En el gráfico es posible apreciar algunas variaciones a lo largo de la serie, así mismo existen algunos cambios respecto a la media de dicha variable ya que se observó una mayor perturbación localizada entre los años 1981 y 1997. Esta perturbación se da en dos contextos históricos, el primero debido a un mal sexenio

¹ y el segundo denominado “El efecto tequila” en 1994 en donde se presentó una devaluación del peso mexicano debido a la falta de reservas internacionales.

Posteriormente existe una mejor estabilidad a partir del año 2001 fecha en donde El Banco de México comienza a establecer la política de objetivos de inflación a 12 meses centrando a la variable en un intervalo de (3% +/- 1%) (Heath J. , 2012, págs. 14, 292, 305).

Por otra parte, puede observarse el grafico de los residuales o rendimientos al cuadrado con la finalidad de observar de mejor forma la localización temporal de las perturbaciones

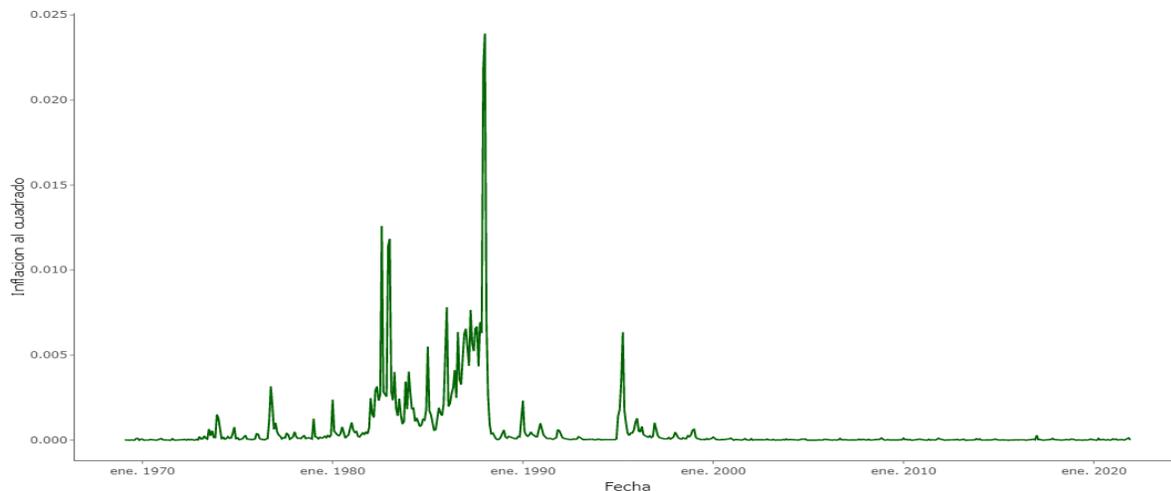
¹ El sexenio de 1983 a 1988, en donde hubo una disminución significativa en el PIB per cápita, el INPC aumentó a un promedio anual de más de 90%, el tipo de cambio pasó de 57 pesos por dólar en promedio en 1982 a 2 mil 284 pesos en 1988 (un aumento de más de cuatro mil por ciento) (Economía UNAM, 2010)

del indicador. Estas perturbaciones se denominan volatilidad. Este es un concepto que puede definirse dentro del contexto de las series temporales como la modificación de la varianza respecto al tiempo (Gujarati & Porter, 2010).

En este sentido, la volatilidad se relaciona con la magnitud de los residuos de una serie en particular sin tomar en consideración el signo de estos, ya sea que sean negativos o positivos, la característica fundamental que comparten estas definiciones es que esa magnitud o variabilidad de la varianza es respecto al tiempo.

Si se operan los datos de la inflación mensual en México, elevando dichos datos al cuadrado se observa lo siguiente en el Gráfico 2.3.

Gráfico 2.3. Inflación mensual al cuadrado: febrero 1969 a diciembre 2021.



Fuente: Elaboración propia utilizando ggplotly() en R.

Como puede notarse, el gráfico anterior incrementa la certeza sobre el tiempo en donde están localizados los acontecimientos históricos de 1960 a 1994 y de 1994 a 2001, que han impactado de forma importante la inflación. Dichos periodos se detallan mejor en las siguientes secciones.

2.2 La Inflación en México: 1960-1994

El sexenio de Gustavo Díaz Ordaz abarcó de 1964 a 1970 y durante este, el país se encontraba en un desarrollo económico que ya venía desde 1954 con el *desarrollo estabilizador*², logrando que el país mantuviera un crecimiento anual del PIB per cápita y la estabilidad de los precios internos o la inflación. Para Tello (2011), el presidente Ordaz apoyó al trabajo agrario, pero también impulsó a la industria manufacturera y las obras de infraestructura de todo el país.

El 1 de diciembre de 1970 Luis Echeverría Álvarez se convierte en el presidente de la república, su objetivo además de continuar con el crecimiento económico era mejorar el descontento social tras el acontecimiento de Tlatelolco en 1968, buscaba beneficiar de ese crecimiento a todos los sectores de la población. Su secretario de hacienda David Ibarra en un análisis para la UNAM, habla sobre el problema que sería proyectar las mismas tasas de crecimiento de los años pasados sin modificar los agentes que provocaban dicho crecimiento. Para Ibarra, era necesario una redistribución del ingreso y mover la demanda de productos internos para que fueran accesibles a la población en general (Whitehead, 1980).

Al comienzo de su mandato en 1971, ante las presiones inflacionarias Echeverría Álvarez opta por una política de austeridad que provoca la escasez de inversión por parte del sector privado. Para finales de este año se registró poco crecimiento económico y mayor inflación y en 1973 se anuncia la ley para regular la inversión extranjera y promover la nacional. Esta medida junto con la intervención que el Estado tenía sobre en la economía provocaron que el sector privado retirara sus inversiones, además se implementara una política de incremento y reducción del gasto público, al final del año si hubo un crecimiento económico, pero la inflación era cada vez mayor (Peña, 1979).

Durante los años siguientes el gobierno federal tomó medidas que respondieran frente a las presiones de la economía. La mejor decisión era la expansión del gasto público, sin

² Época de estabilidad política y económica, progreso desarrollo y orden social en México (Carmona, Montaña, G, Carrion, J, & Aguilar, A, 1970)

embargo, se olvidó el objetivo del crecimiento industrial, el sector agrícola también entró en crisis, no se creaban empleos, había déficit fiscal, la paridad peso dólar dejó de ser fija y la deuda externa aumentó de 6,000 millones de dólares a 20,000 millones de dólares, para agosto de 1976 el peso terminó devaluado 78% (Cantú, 2018).

En el contexto internacional se vivía la *crisis del petróleo*³ y México quería tener una participación más activa y fuerte en el mundo utilizando al hidrocarburo como medio de reconocimiento. El entonces presidente José López Portillo quería recuperar la confianza de los inversores extranjeros para poder pasar de la crisis en la que se encontraba el país, buscaba lograr acuerdos con Estados Unidos que beneficiaran a ambos (Morales, 1986).

En septiembre de 1979 el gobierno de México poseía 5% de las reservas mundiales de petróleo y 3% de las de gas (Smith, 1991), adquiriendo mucha confianza en que el precio del petróleo se mantendría y hasta subiría decidió dejar de lado a los Estados Unidos para concentrarse en Centroamérica, apoyando a los gobiernos en materia política, adquirió créditos que hicieron que la deuda externa aumentara hasta 74.9 millones de dólares (Lusting, 1994).

En 1980 vino el desplome de los precios del petróleo por la sobreoferta y la poca compra debido a la decisión de dejarlo de lado de los países desarrollados, las exportaciones del país se redujeron un 50% (Rico, 2010). La moneda sufrió una gran devaluación pasando de 22 a 80 pesos por dólar, finalmente en su último informe de gobierno, López Portillo se declaró que no podía pagar sus obligaciones financieras y nacionalizó la banca (Arnud, 2006), la inflación anual pasó de ser del 29% en 1981 para llegar a 98.84% en 1982 (BANXICO, 2022).

³ La crisis del petróleo fue una crisis económica originada en la guerra entre Egipto e Israel en 1973, los países miembros de la Organización de Países Árabes Exportadores de Petróleo (OPAEP) subieron el precio del crudo en un 76% y decidieron dejar de exportar gradualmente a los países que hubiesen apoyado a Israel (EUA, Japón y países de Europa), como resultado estos países tuvieron una escasez del producto y empezó una continua alza del precio del petróleo, (Fedesarrollo, 1974).

México estaba en una crisis económica seria, el descontento y la indignación social era profundo debido a la corrupción vista en el sexenio anterior. Luego, Miguel de la Madrid entra como nuevo presidente en 1982 con ideas nuevas y un nuevo plan de acción económica que se basaba mucho en el contexto internacional, el Neoliberalismo, para él se debía limitar la intervención del Estado en el aspecto económico y creía en la autorregulación del libre mercado.

Entre las reformas que se anunciaron para el reordenamiento económico estaba continuar con la reducción del gasto público, bajar la inflación y conservar las inversiones prioritarias, además, garantizó el pago de la deuda externa que en 1983 se renegoció logrando solamente aplazar el pago de los intereses. Las discusiones siguieron durante 1985 y 1986 ya que debido al terremoto que se suscitó no se pudo cumplir con los pagos que se habían pactado, además, el gobierno firmaba acuerdos con el Fondo Monetario Internacional donde se comprometía a detener el aumento salarial, frenar el gasto público e impulsar la apertura comercial (Collado, 2011). Estas decisiones afectaron de manera importante a la clase media y baja del país, además, la inversión pública y privada disminuyeron en un 40%; las importaciones cayeron y para 1984 había un superávit de la cuenta corriente (Collado, 2011).

México enfrentaba también una falta de presencia internacional e intentó reestablecer las negociaciones con Estados Unidos pero el país norteamericano permanecía receloso por los fallos en los acuerdos durante el sexenio pasado. En 1986, México se integró al Acuerdo General sobre Aranceles y Comercio (GATT) para ayudar a la liberación del comercio, sin embargo, con las pocas exportaciones y el aumento de la inflación otra devaluación del peso fue inminente (UNAM, 2002).

Para el final de 1987 el desempleo era masivo y la inflación continuó su crecimiento de manera elevada, su registro anual llegó a 159% (BANXICO, 2022), la Bolsa de Valores Mexicana cayó junto con las del resto del mundo en el llamado "lunes negro", la recesión económica era internacional y el tipo de cambio llegó a 2,298 pesos por dólar (Gómez y Costilla, 2007).

El sexenio siguiente encabezado por Carlos Salinas de Gortari continuó con el trabajo que inició Miguel de la Madrid manteniendo el perfil Neoliberal, los objetivos del libre comercio, la disminución de la inflación y las negociaciones a la deuda externa, UNAM (2002).

En 1989 se dio a conocer la iniciativa Plan Brady elaborado por Nicholas F. Brady, proponía reducir la deuda de los países en desarrollo a través de la “*Recompra de la deuda con descuento y la Conversión de la deuda en instrumentos financieros comerciales*” (Carsten y Gándara, 1990, págs. 303-308). La implementación de este plan redujo en México considerablemente el problema principal de la deuda externa y los intereses, (Medina, 2006). La inflación bajó hasta llegar al 30% anual en 1990 y luego a 18% un año después (BANXICO, 2022).

Luis Medina Peña en su informe escrito para el CIDE en 2006 “Los años de Salinas”, afirma que al disminuir la deuda “*se liberaron recursos para la inversión pública y el gasto social [...], la economía se normalizaba y crecía la confianza en ella empezó a fluir la inversión extranjera y se intensificó la repatriación de capitales mexicanos que habían huido en las épocas de incertidumbre*” (Medina, 2006, pág. 28).

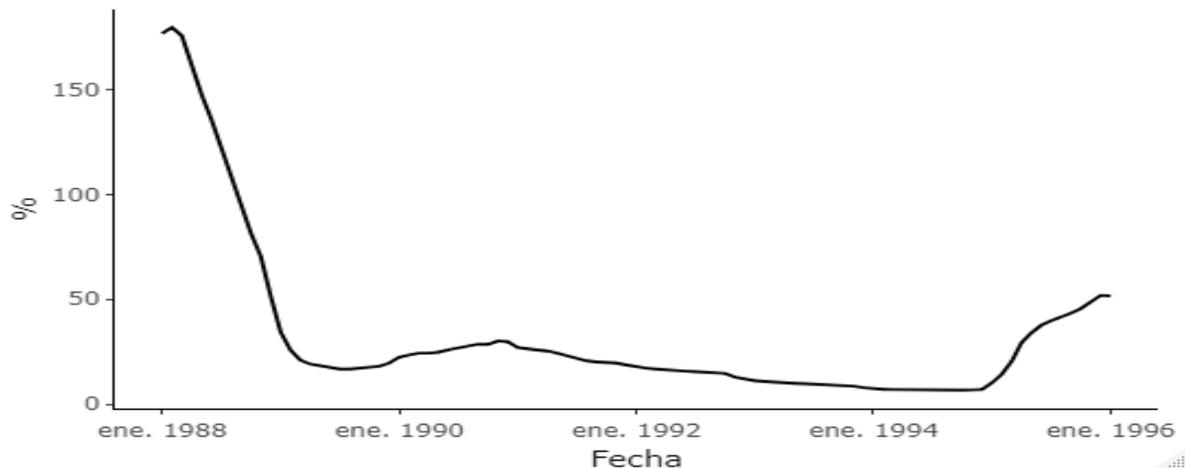
En el nuevo sexenio se continuó con la privatización de las empresas que inició en 1984, el proceso se llevó a cabo con empresas de varios sectores como la industria de transporte (ferrocarriles y aerolíneas) y comunicaciones con la más importante Teléfonos de México, también la banca fue privatizada en 1990. Sin embargo, cuando se adquirieron grandes deudas, el gobierno apoyó para financiarlas, finalmente la mayoría fue vendida a compañías extranjeras. En 1993 se privatizaron 191 empresas paraestatales (Medina, 2006) y la mayoría de las empresas fue liquidada otras se fusionaron, transfirieron o se vendieron al sector privado o social (Sacristán, 2006).

2.3 La Inflación en México: 1994-2022

Para 1995 se había observado una constante disminución de la tasa de inflación y para 1994, había una tasa de inflación del 51.97% considerablemente mayor a la última reportada del 7.1% en 1994. Esto principalmente se debió a dos grandes factores los

cuales fueron el ajuste en el tipo de cambio, así como la depreciación acumulada de la moneda a lo largo de este año, impactando fuertemente el INPC. Este aumento de la tasa de inflación puede observarse en el siguiente gráfico.

Gráfico 2.4. Inflación anual en México de enero 1988 a diciembre 1995.



Fuente: Obtención a través del informe anual de BANXICO de 1995, con datos del INEGI para el INPC y uso de ggplotly() en RStudio

Para 1996 se tuvo una tasa de inflación del 27.7%, valor favorable si se compara con la de 1995. Teniendo en cuenta que a inicios de año se tuvieron las tasas mensuales más altas por los rezagos en los ajustes de los precios que se venían arrastrando desde 1995 y aunado a eso hubo ajustes en las tarifas del metro y servicios telefónicos. Para abril de este año hubo un repunte en la tasa mensual, posterior a esto seguiría con su descenso en el periodo de mayo a agosto, teniendo repuntes en septiembre que son conocidos por el aumento en las colegiaturas para verse aumentada en octubre gracias a la alianza para el crecimiento que impactaría directamente a los precios de las gasolinas y diésel.

En 1997 se pudo observar una tasa de inflación anual mucho menor que la de 1996, siendo del 15.7% anual, la cual se quedó muy cerca de la meta para ese año que fue del 15%. Hablando sobre las variaciones mensuales que se tuvieron a lo largo del año la más alta fue en enero y esto debido a la fluctuación que existió sobre el tipo de cambio, así como los aumentos en los salarios mínimos y otros factores como la alianza del crecimiento la cual entro en vigor en diciembre de 1997.

Entonces, en 1998 la tasa de inflación anual cerró en diciembre con un 18.61% siendo casi tres puntos porcentuales mayor que la del año pasado. Esto fue atribuido a varios factores entre los cuales puede destacarse la inestabilidad en los mercados financieros y la caída del precio del petróleo. La depreciación del peso como era de esperarse impactó en el INPC ocasionando que las expectativas sobre el alza de la tasa de inflación crecieran aumentando el precio de bienes y servicios. Como dato adicional a la inflación de este año se puede destacar que se tuvo esta variación aun con restricciones sobre la política monetaria.

Para 1999 se planteó una meta de una inflación del 13% llegando a esta con una tasa de inflación anual del 12.32% siendo la tasa más baja que se registró durante los últimos 5 años después del incremento abrupto que se presentó en 1995. Esto se debió gracias a que se pudo aplicar una política monetaria enfocada a atacar los efectos inflacionarios ocasionados por factores externos, otro de los factores que también ayudó a la moderación de la inflación en este año fue la apreciación y estabilidad de la moneda nacional.

En el año 2000 nuevamente se pudo alcanzar el objetivo el cual planteaba mantener la tasa de inflación anual por debajo del 10%. Esto se logró con una tasa de inflación anual del 8.96% dejándola como la tasa más baja de los últimos 6 años y muy cerca de la tasa que se tuvo en 1994 la cual llegó a ser del 7.05%.

La correcta aplicación de la política monetaria enfocada en reducir la inflación, así como la estabilidad sobre el tipo de cambio fueron los factores que ayudaron a tener una disminución en la inflación comparada con la de 1999. Este comportamiento se pudo ver desde el mes de enero en el cual se observaba que la variación en el INPC era menor con tendencia a la baja a lo largo del año teniendo una ligera alza en el último trimestre de este.

En 1995 se presenta un alza de los precios siendo una tasa de inflación considerablemente alta ya que se había perdido la confianza sobre la moneda mexicana por estar sobrevalorada. Aunado a esto se aprobó el Fobaproa el cual tenía como intención

absorber las deudas frente a los bancos y así las instituciones financieras pudiesen inyectar más recursos dando pie a la recuperación como se vio a lo largo de su sexenio pasando de una tasa inflacionaria del 51.97% a una del 8.96%.

Para el año 2001, la tasa de inflación anual era del 4.4% siendo esta la menor variación registrada en la historia del INPC hasta el momento. El haber conseguido esta tasa de inflación fue debido a factores como la aplicación de la política de corte preventivo, la desaceleración de la demanda agregada y la estabilidad del precio de la moneda nacional entre otros, siendo este un buen inicio de la presidencia de Fox hablando en términos inflacionarios. Como la tasa de inflación anual más baja en la historia podemos observar su evolución mensual en la Tabla 2.1.

Tabla 2.1. Índice Nacional de Precios al Consumidor (Variación en por ciento)

Mes	Mensual⁴	Acumulada⁵	Anual⁶
Enero	0.55	0.55	8.11
Febrero	-0.07	0.49	7.09
Marzo	0.63	1.12	7.17
Abril	0.5	1.63	7.11
Mayo	0.23	1.87	6.95
Junio	0.24	2.11	6.57
Julio	-0.26	1.84	5.88
Agosto	0.59	2.45	5.93
Septiembre	0.93	3.4	6.14
Octubre	0.45	3.87	5.89
Noviembre	0.38	4.26	5.39
Diciembre	0.14	4.4	4.4

Fuente: Elaboración propia con datos del informe anual de BANXICO para el 2001.

⁴ Variación con respecto al mes anterior

⁵ Variación con respecto a diciembre del año anterior

⁶ Variación con respecto al mismo mes del año anterior

Para el año 2002 a diferencia de los últimos años no se logró llegar a la meta que se tenía la cual era mantener una inflación por debajo del 4.5% y llegando a una tasa del 5.7%. Entre las causas principales a los que se les atribuye el hecho de que no se haya podido llegar a la meta establecida se cita que los precios de la electricidad residencial, así como los del gas doméstico sufrieron aumentos considerables, aunado a esto también se vio afectado por las fuertes cotizaciones en frutas y verduras.

Aun teniendo todos estos precedentes es importante resaltar que la variación de la inflación obtenida contra la que se había propuesto como meta tampoco fue tan abismal gracias a las estrategias de política monetaria. La desaceleración de la demanda, así como un buen comportamiento de los precios en las mercancías pese a la depreciación que sufrió la moneda nacional para ese año. En el año 2003 se tuvo una tasa de inflación anual del 3.98% siendo esta la nueva tasa más baja registrada en la historia, pero aun así quedó por encima de la tasa que se había propuesto como meta para ese año la cual era del 3%.

Los principales factores que se atribuyeron a esta baja inflación fueron los mismos que se habían estado trabajando a lo largo del sexenio de Fox, agregando también la ausencia de perturbaciones importantes en la oferta para el segundo trimestre de este año. Posteriormente en 2004 hubo varias perturbaciones de origen externo las cuales llevaron a la inflación anual a un nivel del 5.19% siendo significativamente mayor a comparación con el del año pasado. Uno de los factores que más afectó al INPC este año aparte de los ya mencionados fue el hecho de que los alimentos procesados también sufrieron alteraciones importantes en sus precios.

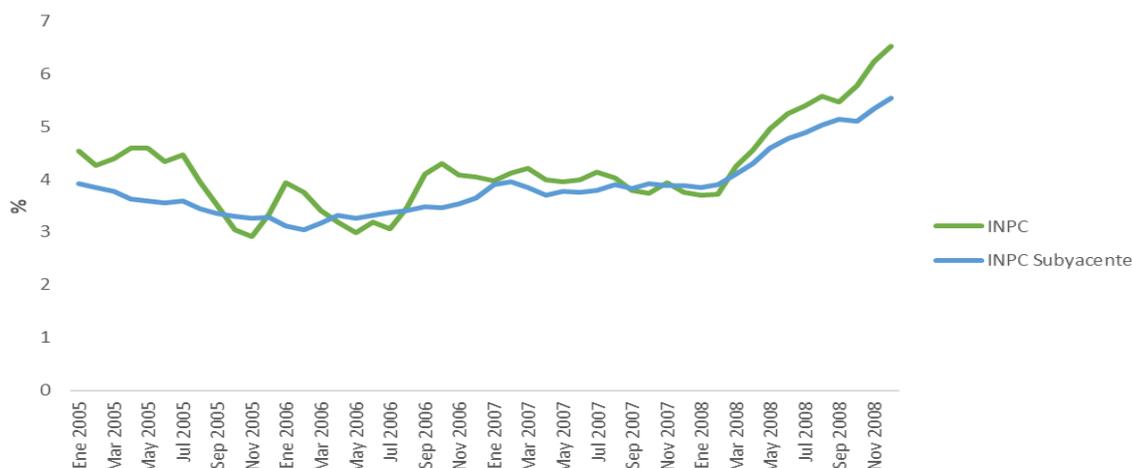
Para 2005 se tuvo una tasa de inflación anual de 3.33% la cual se ve explicada en gran parte a dos factores los cuales fueron la dilución de diversas perturbaciones en la oferta y las acciones de política monetaria que se aplicaron durante el año, lo cual tuvo un resultado favorable logrando tener una variación del INPC menor en comparación al 2004 teniendo como objetivo que la inflación para ese año lograra una convergencia en un 3% anual el cual no se quedó tan alejado.

La inflación para 2006 fue determinada en gran medida gracias al desvanecimiento de perturbaciones sobre la oferta lo cual ocasionó que se tuviese una tasa del 4.05% anual por encima del año pasado. Cabe destacar que durante los primeros meses del año la inflación relativamente fue baja, sin embargo, para los últimos meses de este hubo un incremento significativo en el componente no subyacente del INPC, más particularmente el subíndice de los grupos agropecuarios.

Para 2007 se tuvo una inflación del 3.76% siendo menor que la de 2006 destacando que a inicios de año esta variación se mantenía alrededor de un 4%. La inflación en este año se vio determinada principalmente por tres factores los cuales fueron el aumento en los precios de los alimentos procesados, una inflación no subyacente que logro descender en los últimos meses del año teniendo acciones como el congelamiento de precios en la gasolina, gas y electricidad, así como por último un menor ritmo de crecimiento en los precios de los materiales para la construcción.

En 2008 se presenta la crisis ocasionada por los créditos hipotecarios que llevo a la quiebra a varias instituciones, así como la caída del sistema financiero siendo también esta una de las causas por las que la inflación en este año llegó a un 6.53% anual. Todos los factores externos e internos que se vivían en ese entonces fueron factores fundamentales en el aumento de la tasa de inflación para el año en curso ya que se vieron afectadas las ofertas de bienes y servicios, así como también es importante mencionar que hubo aumentos históricos en los precios internacionales de las materias primas alimentarias, metálicas y energéticas. Además, la moneda nacional sufrió una depreciación importante alrededor del 26%. En el Gráfico 2.5 se puede observar las alteraciones que existieron conforme a la tasa de inflación para este año de crisis:

Gráfico 2.5. INPC vs INPC subyacente⁷ (Variación anual en por ciento).



Fuente: Elaboración propia.

Después de la crisis que se sufrió en 2008, para 2009 la tasa de inflación anual se ubicó en un 3.57% reduciéndola en casi 3 puntos porcentuales en comparación con la inflación del 2008. La reducción del INPC para este año se debió principalmente a cuatro factores los cuales se atribuyen a la reducción de los precios en materiales de construcción, la contingencia sanitaria que se presentó por el brote de influenza contrayendo la demanda del turismo, la eliminación gradual de los choques de oferta que se asocian a las cotizaciones internacionales de los granos, así como el proceso de formación en los precios de los servicios.

Para el cierre del 2010 se tuvo una tasa de inflación anual del 4.4% siendo mayor que la reportada a finales del 2009 (Banco de México, 2011). El haber tenido una tasa más alta al cierre del año se puede explicar gracias a factores como la apreciación cambiaria, incrementos moderados sobre los indicadores salariales y el impacto transitorio sobre la inflación derivado de las modificaciones impositivas y de los aumentos en las tarifas autorizadas por distintos niveles de gobierno (Banco de México, 2011). Un dato adicional es que el descenso de la inflación se vio limitado gracias al incremento que se dio en este periodo al IVA, pasando de 15 a 16% (Banco de México, 2011).

⁷ El conjunto del INPC subyacente contiene a los genéricos con cotizaciones menos volátiles o con evolución más estable, asimismo; mide la propensión inflacionaria de mediano plazo. (INEGI, s.f.)

En 2011 se tuvo un promedio de inflación del 3.41% cerrando el año con una tasa del 3.82% siendo claramente menor a la que se tuvo en 2010 (Banco de México, 2012). Algunos de los factores que mantuvieron esta baja tasa de inflación fueron los atribuidos a la política monetaria que se aplicó y gestionó durante este año tomando en cuenta factores como la intensificación de competencia entre cadenas comerciales, alza de tarifas precios públicos que habían entrado en vigor para 2010 así como la ausencia de presión sobre los precios por parte de los costos laborales (Banco de México, 2012).

Para el 2012 la tasa de inflación al final del año del 3.57% aunque cabe resaltar que la inflación para el resto del año se mantuvo por encima del 4% lo que fue ocasionado por los incrementos de precios en los productos agropecuarios (Banco de México, 2013). Otro dato que es importante destacar para este año es que la variación en la inflación de los precios para las mercancías fue mayor que para el de los servicios, lo cual fue ocasionado por la volatilidad en el tipo de cambio que hubo en este año, así como el alza de los precios internacionales (Banco de México, 2013).

Para el 2013 la tasa de inflación anual cerró en un 3.97% quedando casi un punto porcentual del objetivo de convergencia, destacando que a finales de este año existieron ajustes no anticipados sobre los precios en el transporte público en ciertas regiones del país, así como el aumento de los precios en ciertos productos agrícolas lo que llevó a un repunte de la tasa en los meses de noviembre y diciembre respectivamente (Banco de México, 2014).

Para el 2014 la tasa cerró en un 4.08% quedando por encima del margen que se tenía de más menos un punto porcentual alrededor del 3%. Para este año hubo un importante deterioro originado por una caída significativa en el precio del petróleo, así como una apreciación en el dólar dando lugar a una alta volatilidad sobre los mercados financieros y puntos de vulnerabilidad financiera en economías emergentes como la de México (Banco de México, 2015).

Hacia el 2015 se presenta un récord en cuanto a la variación mínima histórica reportada desde que se empezó a medir el INPC siendo que para diciembre de 2015 se registró una tasa de inflación anual del 2.13% (Banco de México, 2016). Es importante mencionar que las circunstancias en la que la tasa de inflación llegó a ser muy baja fueron complejas ya que se presentaban factores como la debilidad de la economía mexicana frente al crecimiento internacional teniendo bajas expectativas de crecimiento, así como la constante caída del precio del petróleo (Banco de México, 2016).

En 2016 se cerró el año con una tasa de inflación anual del 3.36%, tomando en consideración que durante este año se conservó un ambiente de alta volatilidad a consecuencia de factores externos que, en este caso, se atribuye principalmente a la incertidumbre sobre el proceso de normalización de la postura monetaria de los Estados Unidos de América y sus procesos electorales en donde ganaría Donald Trump para asumir el cargo en 2017 (Banco de México, 2017).

Para 2017 hubo una considerable inflación, la cual para diciembre alcanzó un 6.77%, este efecto se debió a varios factores que tuvieron gran impacto tanto externos como internos (Banco de México, 2019). Los factores que concretamente afectaron a la tasa de inflación en este año, uno de ellos fueron el aumento en los precios de los energéticos en concreto gasolina y gas LP, aumentos en las tarifas de transporte público, así como aumento en los precios de ciertos productos agropecuarios (Banco de México, 2018).

Para diciembre de 2018 se reportó una tasa de inflación anual del 4.83%, siendo considerablemente menor a la que se tuvo el año pasado tomando en cuenta que para este periodo existió una fuerte restricción del financiamiento externo (Banco de México, 2019). Otros factores que afectaron directamente a la inflación de este año han sido atribuidos a la depreciación que ha tenido el petróleo desde el 2014, así como a la incertidumbre que existía en ese momento sobre la relación comercial con Estados Unidos y Canadá (Banco de México, 2019).

En 2019 se reportó una tasa del 2.83% anual, manteniéndose en el margen de convergencia objetivo que ha sido una tasa del 3% más menos un punto porcentual (Banco

de México, 2020). Cabe destacar que entre los factores que afectaron a la inflación en este año, se tiene que el peso mexicano sufrió una apreciación y aunado a esto se mantuvo el estancamiento sobre el crecimiento económico durante todo el año (Banco de México, 2020).

Para 2020 la tasa de inflación anual en diciembre se reportó en un 3.15% manteniéndose en el margen de la meta de convergencia en un 3% pese a que existía bastante incertidumbre ocasionada por la pandemia del COVID-19 la cual detuvo la economía a nivel mundial por el confinamiento (Banco de México, 2021).

La volatilidad que se presentó durante este periodo afectó significativamente a la situación financiera del país y no solo al país si no que a nivel mundial hubo afectaciones por posibles rebrotes de esta pandemia frenando la recuperación de las economías (Banco de México, 2021). Si se toma en cuenta todo el contexto que esta pandemia ocasionó a nivel económico se es de esperar que en los próximos años exista una recesión a nivel mundial, así como tasas de inflación más altas de las que se venían manejando (Banco de México, 2021).

Para el cierre de 2021 se reportó una tasa del 7.36% siendo esta la más alta registrada en las últimas dos décadas a causa de las secuelas financieras que dejó la COVID-19 (Banco de México, 2022). Como respuesta a estos riesgos inflacionarios BANXICO ha venido aplicando una política monetaria con la intención de poder evitar afectaciones en las expectativas sobre los niveles inflacionarios (Banco de México, 2022). Hablando de factores externos y de la economía global los demás países han ido recuperándose de la secuela que ha dejado el COVID-19 dejando un panorama para 2022 poco favorable en el que se estima que el crecimiento de la economía será a la baja (Banco de México, 2022).

Para el 2022 que es el año en curso a la elaboración de este trabajo la tasa de inflación se ha visto afectada por varios factores y variables macroeconómicas, entre ellas podemos seguir percibiendo la gran crisis financiera que se ocasionó por la pandemia del COVID-19, así como los conflictos que han existido entre Ucrania y Rusia. Este conflicto ha afectado el comercio internacional por lo que se sigue pronosticando una muy posible

estanflación, sabiendo que el último dato reportado de la tasa de inflación anual fue del 8.41% en el mes de octubre.

2.4 Factores asociados a la Inflación en México

Con la revisión de la historia de la inflación en México se identifican diferentes factores y causas que en su conjunto provocaron el aumento generalizado de los precios. Entonces como ahora, el contexto y mercado internacional afectan de manera directa a la economía de nuestro país, las relaciones comerciales de México con los demás países y principalmente con Estados Unidos hacen que inevitablemente se mueva con la economía mundial. Igual que sucedió con las crisis económicas del petróleo, la caída de la bolsa y como actualmente sucede con la guerra entre Rusia y Ucrania, y la especulación alrededor del conflicto o algunos otros sucesos políticos que provocan la inestabilidad en los tipos de cambio y la volatilidad en la bolsa de valores.

Un hecho sin precedentes que afectó al mundo entero y a su estabilidad económica fue la crisis sanitaria que comenzó en marzo del 2020, un golpe duro que en su momento tuvo diversas repercusiones debido a las medidas que fueron tomadas para evitar la propagación del virus y otras que se están viviendo actualmente con la recuperación después de la pandemia.

El crecimiento económico de un país es sin duda un factor importante para la inflación esperada, el Producto Interno Bruto es el indicador por el cual se mide este rubro, por el lado del gasto sabemos que se involucra al consumo de los hogares, la inversión, el gasto público, y las exportaciones menos las importaciones, con la pandemia muchos países cerraron sus fronteras impidiendo el paso de los productos internacionales, el comercio se detuvo y la caída de todas las exportaciones fue inevitable, como resultado el crecimiento económico cayó a -8.20% anual y la pobreza aumentó respecto a 2019, además durante la pandemia se perdieron en México 125 millones de empleos, la mayoría del sector informal y el de servicios, además de que se incrementó la desigualdad entre la población debido a las limitaciones tecnológicas para el trabajo o la escuela en línea. (CONSAR, 2021)

El Banco de México en su reporte trimestral de junio del 2021 habla sobre el repunte inevitable de la inflación para Estados Unidos después de la baja experimentada en el nivel de precios del año anterior, la recuperación del COVID trajo un alza de la inflación que fue mayor a la esperada (BANXICO, Factores Asociados a la Evolución Reciente de la Inflación en Estados Unidos, 2021). En México para el segundo trimestre del 2021 el aumento se debió al incremento en la gasolina y posteriormente en los energéticos que se han visto limitados por el precio de los agropecuarios y el gas LP, además se previó que derivada de la misma recuperación la inflación estaría arriba del 5% para todo el 2022 debido a la escases de los insumos y los problemas para la transportación y el almacenaje de los productos, esperando que en el inicio del 2023 se pueda llegar a la meta del 3% (BANXICO, Informe trimestral Abril- Junio 2021, Resumen visual, 2021).

Expuesto el panorama sobre la evolución de la inflación en México, en el capítulo siguiente se presenta la metodología de investigación mediante la cual se logra la prueba de hipótesis de la tesis.

Capítulo 3. Metodología de la investigación

La metodología de esta tesis se presenta en 3 apartados principales: 1) Fuentes de datos, 2) Técnicas de investigación, 3) Variables de investigación. A continuación, se presenta cada uno de estos apartados.

3.1 Fuente de datos

El enfoque analítico de esta tesis se basa en el de las series temporales dadas las características de los datos estudiados. Para el desarrollo de los análisis de este proyecto fue necesario contar con información de cinco indicadores: Índice Nacional de Precios al Consumidor, Expectativas inflacionarias, Datos de la oferta y demanda agregada, El tipo de cambio al cierre de jornada, así como precio de venta máximo y mínimo diario de la paridad peso/dólar, y el Salario Mínimo.

Respectivamente, cada uno de los indicadores anteriores se obtuvieron de distintas fuentes. Los datos históricos del Índice Nacional de Precios al Consumidor se obtuvieron del portal de internet del Instituto Nacional de Estadística y Geografía (INEGI), los datos recopilados tienen el año base 2018.

Las expectativas inflacionarias se extraen del informe mensual de la encuesta sobre las expectativas de los especialistas en economía del sector privado realizadas por el Banco de México a través del Sistema de Información Económica en el portal de BANXICO.

Los datos históricos de la oferta y demanda agregada se encuentran disponibles en el Sistema de Información Económica en el portal de BANXICO. Por otra parte, el tipo de cambio al cierre de jornada, así como precio de venta máximo y mínimo diario de la paridad peso/dólar se consulta en la página del Banco de México. Mientras que la consulta del salario mínimo en México está disponible en el portal del Gobierno de México en la tabla de Salarios Mínimos Generales y Profesionales por Áreas Geográficas proporcionada por la Comisión Nacional de Salarios Mínimos (CONASAMI).

3.2 Variables de investigación

Para implementar los métodos de pronóstico propuestos en esta tesis es necesario definir a las variables que serán utilizadas como sigue: la variable dependiente (INPC) la cual será pronosticada a futuro, y las variables exógenas o independientes son aquellas que tienen relación con la variable principal (dependiente) con sustento en la teoría o la estadística y que pueden afectar o explicar su comportamiento, en nuestro caso son las siguientes:

- 1) La inflación que será medida a través del Índice Nacional de Precios al Consumidor como variable dependiente, la serie tiene temporalidad mensual y se tomarán en cuenta los datos a partir de 1969 para los modelos SARIMA, GARCH y de Redes Neuronales, para el modelo SARIMAX se tomará en cuenta la temporalidad a partir del año 2000.
- 2) La encuesta realizada a expertos proporciona información respecto a las expectativas inflacionarias, se encuentran disponibles a partir del año 1999 y la periodicidad de los datos es mensual.
- 3) La oferta y demanda agregadas se miden en millones de pesos, sin embargo, los datos están disponibles a partir de 1994 y la periodicidad es trimestral por lo que se realiza un ajuste a los datos para que exista coincidencia en el tema de la temporalidad.
- 4) El tipo de cambio es una recopilación de datos diarios desde 1992, el cual representa el valor del peso mexicano frente al dólar, de igual modo se implementó un ajuste para que exista coincidencia en la temporalidad mensual que en este trabajo se toma en cuenta.
- 5) El salario mínimo mensual presentada desde 1992, el valor del salario mínimo durante cada mes.

3.3 Técnica de investigación

Las series temporales se han caracterizado por ser una herramienta estadística que nos permiten conocer un posible valor futuro de una determinada variable de estudio basado en su comportamiento pasado a través del tiempo. Existe una gran variedad de modelos matemáticos de series de tiempo que permiten pronosticar la inflación. Es importante

señalar que cada modelo tiene características diferentes y su empleabilidad va a variar acorde con los datos disponibles para la investigación. Algunos ejemplos de estos modelos son los siguientes:

El modelo lineal generalizado, que explica a una variable dependiente en función de otra variable con la que guarde una relación más un término de error, que no es otra cosa más que la diferencia entre el valor real observado y el pronosticado por el modelo, este modelo puede ser univariante o multivariante. La ecuación que representa a este modelo sería:

$$y = \beta_0 + \beta X + \varepsilon$$

En donde: y : es la variable dependiente; β_0 : el intercepto de la recta; β : representa la pendiente; X : es la variable independiente; ε : el término de error.

Este modelo puede generalizarse a

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Con n variables explicativas de y .

Por otra parte, se desarrolla la metodología Box – Jenkins, la cual fue publicada en 1970 por estos autores para pronosticar series de tiempo¹³. Es un proceso que consta de cuatro etapas (que se verán a profundidad más adelante), de esta metodología también conocida como metodología ARIMA (Delgadillo, Ramírez , leos, Salas, & Valdez, 2016) se desprenden las ecuaciones que para explicar de mejor manera se deben conocer los conceptos de:

- *Ruido blanco* será el componente aleatorio que tiene media cero y varianza constante (Hernández, 2015).
- *Caminata aleatoria* “el valor de mañana de una serie es el valor de hoy más un cambio impredecible” (Massó & Pons, 2018), la caminata de la variable que se vaya a explicar sigue pasos aleatorios.

Otro concepto sumamente importante es la estacionariedad, ya que este determina la ecuación o modelo que será utilizado para el análisis del conjunto de datos. Una serie será

¹³ En su libro Forecasting and Control publicado originalmente en 1970*

estacionaria en media cuando su media y varianza sean “*estables o constantes en el tiempo*” (Gujarati & Porter, 2010), que no exista ninguna tendencia de los datos hacia crecer o decrecer. Si una serie de tiempo sigue una caminata aleatoria es no estacionaria (Massó & Pons, 2018).

Con los conceptos anteriores es posible definir entonces los modelos en los que tendrá sustento el presente trabajo, no todos serán aplicados, pero es necesario introducirlos para un mejor entendimiento de los procedimientos analíticos desarrollados.

3.3.1 Modelo ARIMA (p,d,q)

a) Modelo autorregresivo AR(p)

Un proceso autorregresivo puede definirse como un modelo que permite pronosticar un valor futuro donde la variable dependiente (y) tiene dependencia con su valor rezagado en el periodo p más un término aleatorio (Gujarati & Porter, 2010).

La ecuación para un proceso con media $\mu \neq 0$ se define como

$$(Y_t - \mu) = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + \varepsilon_t$$

Donde: $\phi_i \forall i = 1, 2, \dots, p$: es la proporción del valor que toma y en el tiempo; μ : es la media de Y ; ε_t que es el error aleatorio (ruido blanco). De manera general, $\phi(B)Y_t = \mu + \varepsilon_t$, y usando el operador de retardo B : $\phi(B) = 1 - \phi_1B - \phi_2B^2 + \dots + \phi_pB^p$.

b) Modelo de medias móviles MA(q)

El modelo de medias móviles por su parte sería donde la variable y depende de un promedio móvil de sus términos de error pasados y presentes más una constante (Gujarati & Porter, 2010), y su ecuación es la siguiente:

$$Y_t = c + \theta_1\varepsilon_t + \theta_2\varepsilon_{t-1} + \theta_3\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q}$$

Donde: $\theta_i \forall i = 1, 2, \dots, p$: es la proporción del valor que toma y en el tiempo; c : es un término constante; ε_t : es el error aleatorio (ruido blanco).}

O de manera general $Y_t = \mu + \theta(B)\varepsilon_t$. Usando el operador de retardo B de media móvil $\theta(B) = 1 + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q$.

De los dos modelos anteriores, se deriva el proceso auto regresivo de medias móviles ARMA(p,q) donde la variable dependiente y tendrá p términos autorregresivos y q términos de promedios móviles. La ecuación del modelo sería:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$\phi(B)Y_t = \theta(B)\varepsilon_t$$

Un modelo es ARMA cuando el proceso es estrictamente estacionario en media, sin embargo, cuando una serie de tiempo no cumple con esta característica se le pueden aplicar primeras diferencias a la variable dependiente y hasta n diferencias para cumplir con la estacionariedad, sin embargo no es sugerente pasar de las primeras 2 diferencias ya que al hacer procesos de diferenciación más prolongados ($d > 2$) se corre el riesgo de pérdida de información que afecta significativamente a los pronósticos. Cuando estas diferencias son aplicadas la ecuación que debe ajustarse es un proceso ARIMA(p,d,q) ya que el proceso pasó a ser autorregresivo integrado de medias móviles en donde d es el número de veces que se tuvieron que aplicar las diferencias. El proceso se puede escribir como:

$$(1 - \phi_1 B + \dots - \phi_p B^p)(1 - B)^d Y_t = \mu + (1 - \theta_1 B - \theta_q B^q)\varepsilon_t$$

Si se utiliza un operador de diferencia $\nabla = 1 - B$ la ecuación queda

$$\phi_p(B)\nabla^d Y_t = \mu + \theta_q(B)\varepsilon_t$$

Donde: $\phi_p(B)$: es el polinomio autorregresivo; $\theta_q(B)$: representa el polinomio de medias móviles; ∇^d : es el operador de diferencias.

Un modelo ARIMA debe cumplir los supuestos de:

- 1) Estacionariedad en media $E(Y_t) = E(Y_{t+m})$, en varianza $E(Y_t - \mu)^2 = E(Y_{t+m} - \mu)^2$ y una función de correlación constante a través del tiempo (Massó & Pons, 2018).
- 2) Invertibilidad del modelo: las proporciones o pesos de las observaciones pasadas Y_{t-1} disminuyen con el tiempo, esto se cumple cuando $\theta < 1$ (Massó & Pons, 2018).

- 3) Los residuos de ruido blanco tienen media igual a 0, varianza constante y no hay autocorrelación (Massó & Pons, 2018).

Una serie de tiempo ARIMA además puede presentar un componente *estacional* que en términos generales significa que está siguiendo un patrón cada cierto (pero siempre en el mismo) periodo de tiempo, y este patrón de repetición podría ser explicado por diferentes factores externos sociales, políticos o económicos.

A la ecuación conocida se le agrega el término estacional y se crea la ecuación SARIMA o simplemente ARIMA estacional (p,d,q) (P,D,Q) los parámetros P,D,Q se van a encargar de modelar el componente estacional, entonces la ecuación quedaría como:

$$\Phi_P(B^S)\phi_p(B)\nabla_S^D\nabla^dY_t = \mu + \theta_q(B)\Theta_Q(B^S)\varepsilon_t$$

Donde: ϕ_p : representa el orden del polinomio autorregresivo regular; θ_q : el orden del polinomio de medias móviles regular; $\Phi_P(B^S)$: es el orden del polinomio autorregresivo estacional; $\Theta_Q(B^S)$: representa el orden del polinomio de medias móviles estacional; ∇_S^D : son las diferencias estacionales; ∇^d : las diferencias regulares.

3.3.2 Modelo ARIMAX

Un modelo ARMAX(p,q) es una generalización de un ARMA con la característica de agregar una o más variables exógenas relacionadas al modelo con una respuesta estacionaria (Romero, 2019). La forma general de este modelo con su operador de retardo sería:

$$\phi(B)Y_t = c + x'_t\beta + \theta_q(B)\varepsilon_t$$

Donde: x'_t : es el vector con los valores de las t variables predictoras exógenas; β : el coeficiente asociado a cada variable exógena que trata de explicar al modelo.

El modelo ARIMAX (p,d,q) queda de la forma siguiente:

$$\Phi_P(B)\nabla^dY_t = \mu + x'_t\beta + \theta_q(B)\varepsilon_t$$

A la ecuación ARIMAX también se le puede agregar el componente estacional para obtener un modelo ARIMAX (p,d,q) (P,D,Q) (Romero, 2019), y queda de la siguiente forma:

$$\Phi_P(B^S)\phi_p(B)\nabla_s^D\nabla^dY_t = \sum_{I=1}^n x_{it}\beta + \theta_q(B)\Theta_Q(B^S)\varepsilon_t$$

En donde la única diferencia con el modelo SARIMA(p,d,q) es que ahora tendremos variables exógenas y el operador de diferencias ha de afectar de igual modo a dichas variables.

3.3.3 Modelos ARCH y GARCH

Si una serie de tiempo tiene problemas para ajustar una ecuación del tipo ARIMA aun cuando se le aplicaron diferencias al proceso puede ser debido a un problema de la varianza, aunque parezca que la serie sea estacionaria, la parte de la varianza podría no estarse tratando de forma adecuada, esto se presenta comúnmente en las series de tiempo financieras, por ejemplo, los precios de las acciones, tasas de inflación y en general algunos índices de carácter económico que suelen presentar un comportamiento volátil.

En un contexto de series temporales, si nos centramos en los residuos o en los rendimientos los cuales van a variar respecto al tiempo, en este caso llámese ε_t , una manera eficiente de medir la volatilidad es obtener los residuos o rendimientos al cuadrado ε_t^2 , al hacer esta transformación podemos enfocarnos en resolver el problema de los signos, es decir, solo nos centramos en la magnitud de los residuos y si los observamos en un gráfico, podemos apreciar de una mejor manera a la volatilidad que se encuentra implícita en la serie.

Para modelar de una mejor manera este tipo de serie económica se tiene la necesidad de controlar aquellos periodos con una mayor volatilidad, para ello se introducen los Modelos ARCH y GARCH a continuación.

3.3.3.1 Modelos ARCH

Dentro del análisis de las series de tiempo, el enfoque más común para comprender de mejor manera a la volatilidad es utilizando los Modelos Autorregresivos con Heterocedasticidad Condicional (ARCH) propuestos por Robert Engle (1982). Estos modelos están compuestos de dos ecuaciones, una para modelar la media de la serie la cual puede modelarse con la metodología ARIMA y otra para modelar la varianza. Estos modelos usualmente pueden definirse como modelos no lineales donde se asume que los errores se distribuyen de forma normal con media cero, pero en este caso en particular la varianza está condicionada a los errores pasados o a su comportamiento anterior, técnicamente es el valor esperado del término de error al cuadrado condicionado a los valores anteriores.

Para introducir una definición más formal de los modelos ARCH, puede partirse del modelo más sencillo dentro del campo de las series de tiempo, este modelo será un proceso autorregresivo que se define como:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

En donde: y_t representa el valor que toma la serie en un momento particular t ; ϕ_0 es la constante o pendiente de la serie; ϕ_1 es el coeficiente asociado al rezago anterior de la serie; y_{t-1} representa al rezago inmediato anterior de la serie de tiempo; ε_t es el error asociado a la serie.

De acuerdo con la metodología ARMA y/o ARIMA el término de error ε_t debería tener una distribución en particular para caracterizarse como ruido blanco, es decir la distribución asociada a los errores es una distribución Normal con media cero y una varianza constante, expresado en términos matemáticos:

$$\varepsilon_t \sim N(0, \sigma^2)$$

A la hora de analizar variables de índole económico, es posible notar que dichas series tienden a ser volátiles, por lo que seguir considerando este supuesto sería algo incorrecto. Como ya se mencionó con anterioridad para un modelo ARCH los errores dependen o en otras palabras están condicionados a sus valores pasados, ahora bien, al desglosar esta

idea y juntarla con la característica heterocedástica que por naturaleza presentan las series económicas, se obtiene la siguiente expresión:

$$\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q}$$

Como la volatilidad es sinónimo de varianza, puede obtenerse dicha medida de dispersión como sigue:

$$\sigma^2 = \text{Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q})$$

Al conjuntar este supuesto con la idea de que los errores siguen teniendo una media cero y alterando la parte de la varianza constante, se desarrolla la idea principal de este tipo de modelos como sigue:

$$\sigma^2 = \text{Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q}) = E[(\epsilon_t - E(\epsilon_t))^2 | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q}]$$

Como $E(\epsilon_t)$ es igual a cero, la expresión se simplifica como sigue:

$$\sigma^2 = \text{Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q}) = E[\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-2}, \epsilon_{t-3}, \dots, \epsilon_{t-q}]$$

La expresión anterior confirma que la volatilidad de la serie está condicionada a los errores pasados o a su comportamiento anterior, dicho de otro modo, la varianza condicional de la serie la cual sigue conservando el supuesto de media igual a cero en los errores, es igual a la varianza condicional de los residuos al cuadrado. Ahora bien, el modelo más simple en términos de volatilidad es la ecuación ARCH (1), la cual tiene la siguiente estructura:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$$

En donde: σ_t^2 Representa a la varianza de la serie en el periodo t ; α_0 Es un término constante del modelo; α_1 Es el coeficiente que acompaña a los errores del periodo anterior; ϵ_{t-1}^2 Son los errores al cuadrado anteriores inmediatos de la serie.

Al incluir los errores pasados de la serie se produce el efecto autorregresivo y con él se tiene más certeza de explicar a la serie por sí misma con los efectos anteriores que haya presentado. Así mismo este modelo puede generalizarse a una ecuación ARCH(q), la

única diferencia es que en este modelo incluimos más rezagos anteriores de los errores del modelo, la ecuación de este proceso se define como sigue:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2$$

Y en su forma compacta puede escribirse como:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

Con la explicación anterior y retomando el ejemplo del modelo AR (1) se tiene lo siguiente:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

En donde ahora, el término de error tendrá una distribución normal con media cero y una varianza que se explica con el modelo ARCH, es decir:

$$\varepsilon_t \sim N(0, \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2)$$

a) Condiciones de los modelos ARCH(q)

Poder justificar que existen efectos ARCH en la serie: Para ello podemos realizar pruebas de hipótesis para confirmar que cada uno de los **q** rezagos de la serie de tiempo en cuestión son significativos, en términos más sencillos confirmar que son diferentes de 0.

La suma de los parámetros debe ser menor a 1: Este primer punto debe satisfacerse ya que de lo contrario el modelo se volvería inestable y por el contrario no obtendríamos un buen ajuste del modelo.

No negatividad: Como el punto principal del supuesto es explicar la serie con los errores al cuadrado o en otras palabras como se está trabajando con la varianza condicional de la serie, es lógico afirmar que todos los coeficientes y/o constantes asociados al modelo ARCH (1) deben ser positivos, es decir: $\alpha_0 \geq 0$ y $\alpha_1 \geq 0$

Y para un modelo general, se cumple que:

$$\alpha_i \geq 0 \forall i = 1, 2, \dots, q$$

b) Limitaciones de los modelos ARCH

Usualmente para identificar el orden de la ecuación ARCH hacemos uso de la función de autocorrelación parcial (PACF) observando aquellos rezagos que son significativos, sin embargo hay ocasiones en donde existen demasiados rezagos para considerarlos a todos y esto nos llevaría a tener un orden ARCH bastante alto ergo algunos coeficientes de la ecuación pueden volverse negativos haciendo difícil una interpretación, así mismo el modelo ya no estaría respetando el supuesto de parsimonia haciendo que las predicciones no sean del todo correctas.

3.3.3.2 Modelos GARCH

Para introducir este tipo de modelos pueden plantearse algunas limitaciones que tiene el modelo ARCH por sí mismo, que, como ya se mencionó en algunas ocasiones la misma naturaleza de la serie provoca que se puedan considerar demasiados rezagos a la hora de ajustar un modelo, por esta razón es que nacen los Modelos Generalizados Autorregresivos Condicionales Heterocedásticos (GARCH) que fueron propuestos por Bollerslev y Taylor (1986).

Estos modelos pueden considerarse como una extensión de los modelos ARCH en los cuales, aparte de considerar los errores pasados de la serie, también incluiremos el efecto de la varianza condicional rezagada q periodos. Más formalmente, el modelo más simple para esta familia sería un modelo GARCH (1,1) el cuál presenta la siguiente estructura:

$$\sigma_t^2 = \omega + \phi_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

En donde ahora: ω : representa una constante del modelo; β_1 : que representa al coeficiente que acompaña a la varianza condicional rezagada 1 periodo; σ_{t-1}^2 : resulta ser la varianza condicional rezagada 1 periodo.

Este proceso de igual forma puede generalizarse a un modelo GARCH(p,q), el cual se representa como sigue incluyendo los p rezagos para los errores así como también los q rezagos para la varianza condicional de periodos anteriores:

$$\sigma_t^2 = \omega + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \varepsilon_{t-2}^2 + \dots + \phi_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_q \sigma_{t-q}^2$$

En su forma más compacta, la ecuación puede escribirse como:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \phi_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2$$

Existe una demostración matemática que permite representar el modelo GARCH como si fuese un modelo ARMA, cabe señalar que dicha demostración sigue siendo en torno de predicción de la varianza del modelo, para confirmar lo anterior la demostración queda como sigue:

- 1) Recuperando el concepto de error, el cual puede definirse como la diferencia entre el valor observado y valor que ajustamos, ahora bien, en relación con la varianza de un modelo GARCH, este error puede expresarse como $\varepsilon_t = \varepsilon_t^2 - \sigma_t^2$.
- 2) Dado que la variable de interés es σ_t^2 esta se despeja de la ecuación anterior y conociendo ahora su valor, puede sustituirse en la expresión resultante en la ecuación del modelo GARCH (1,1), esto es $\sigma_t^2 = \varepsilon_t^2 - \varepsilon_t$; $\varepsilon_t^2 - \varepsilon_t = \omega + \phi_1 \varepsilon_{t-1}^2 + \beta_1 (\varepsilon_{t-1}^2 - \varepsilon_{t-1})$.
- 3) Como paso siguiente puede ordenarse el término de error al cuadrado dejándolo en el lado izquierdo de la ecuación, así mismo también puede obtenerse el factor común y ordenar la ecuación, teniendo:

$$\varepsilon_t^2 = \omega + \phi_1 \varepsilon_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2 - \beta_1 \varepsilon_{t-1} + \varepsilon_t; \varepsilon_t^2 = \omega + \varepsilon_{t-1}^2 (\phi_1 + \beta_1) - \beta_1 \varepsilon_{t-1} + \varepsilon_t$$

El resultado que proporciona esta demostración es una forma alternativa de representar una ecuación GARCH (1,1) como si fuese un proceso ARMA (1,1) con la distinción de que este proceso ARMA se realiza sobre los errores al cuadrado. El motivo principal para usar este tipo de modelo radica en que nos ayuda a evitar modelos que contengan demasiados parámetros ergo evitamos el sobreajuste de la serie al modelar la varianza y evitamos que los coeficientes asociados al modelo violen el supuesto de no negatividad. El hecho de incluir a la varianza condicional de periodos anteriores nos permite tener un mejor control sobre la volatilidad, por lo que en la mayoría de las veces los investigadores prefieren

utilizar los modelos GARCH para explicar de mejor manera a la volatilidad de las series de tiempo.

Un problema que suelen presentar estos modelos es que no responden de manera adecuada a variaciones negativas o positivas, es decir, a cuando el precio de una acción sube o bajan, lo toman de forma simétrica y entonces los pronósticos pueden verse afectados y no resultar tan precisos.

3.3.4 Modelo de Redes neuronales

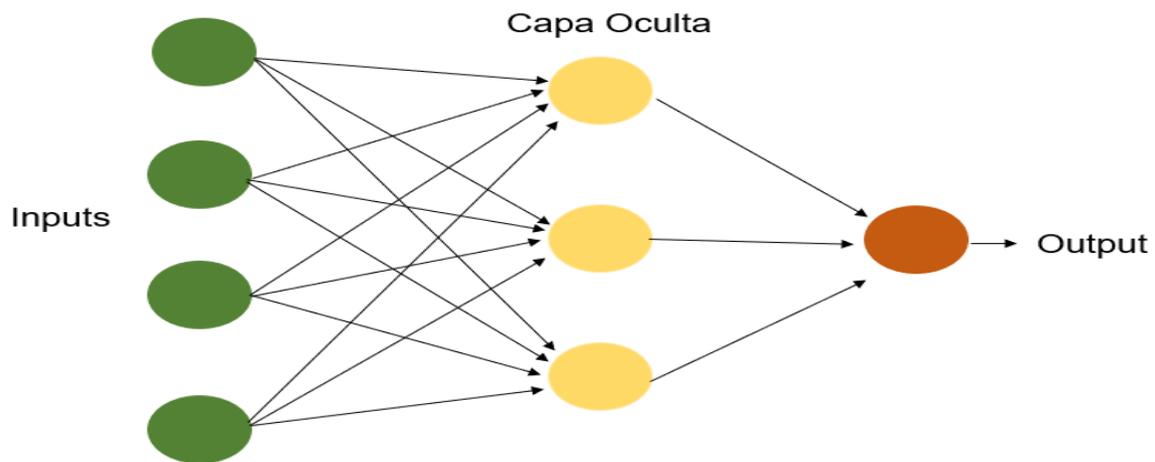
Acorde a la definición propuesta por Rob Hyndman en su libro: “Forecasting: Principles and Practice”, los modelos de redes neuronales están basados en una interpretación matemática del cerebro o mejor dicho de las neuronas en un sentido biológico que permiten el aprendizaje. Este tipo de modelos permite manejar relaciones no lineales complejas entre la variable dependiente y los predictores.

Una red neuronal puede verse como un conjunto de neuronas, que, en este ámbito, pueden definirse como una función no lineal que se encuentra acotada y parametrizada¹⁴. La arquitectura de una red neuronal puede ser variada, sin embargo, para fines de esta tesis hemos de prestar especial atención a aquellas denominadas multilayer feed-forward network o red neuronal multicapa alimentada hacia adelante.

La estructura de este tipo de redes puede apreciarse mejor en el Gráfico 3.1 en donde podemos apreciar que se tiene una capa de entrada con n número de neuronas denominadas Inputs (que en este ejemplo suman un total de 4 neuronas), seguidas por una capa intermedia la cual es denominada como capa oculta (conformada por 3 neuronas) y por último una sola neurona como salida de este proceso, comúnmente denominada Output.

¹⁴ Dreyfus, G. (s.f). Neural Networks: Definitions and Properties. En Neural Networks: Methodology and Applications (págs. 2-6).

Gráfico 3.1. Representación gráfica de una red neuronal alimentada hacia adelante.



Fuente: Elaboración propia con la información contenida en Forecasting: Principles and Practice de Rob. Hyndman (<https://otexts.com/fpp3/nnetar.html>)

De esta forma cabe resaltar que la capa de entrada es aquella que recibe toda la información de inicio para después ir alimentando las capas consecuentes para así al último pueda ser expulsado el valor esperado en la capa de salida, por lo que, si una pequeña parte del sistema ha sido dañada, esto no afectará al funcionamiento total de la misma (Vidal Gonzalez, 2015).

El funcionamiento de la red en su conjunto comienza definiendo aquellos valores que han de fungir como Inputs, usualmente en la capa de entrada los valores de las conexiones son fijas, estos valores son ponderados por medio de una combinación lineal, esta ponderación se hace a través de pesos (w_i), los cuales son determinados en función a los datos de entrada y son utilizados para prevenir que esta información sea demasiado larga¹⁵, así mismo dichos datos son sumados a un sesgo o bias (b_i) que al igual que los pesos, estos son estimados de los datos disponibles para el pronóstico. Expresado matemáticamente se tiene la siguiente combinación lineal:

$$y_i = \sum_{i=1}^n w_i * x_i + b_i$$

¹⁵ Esta restricción se conoce como parámetro de decaimiento y usualmente se establece como un valor fijo de 0.1 para prevenir el alargamiento de los datos de entrada. (<https://otexts.com/fpp3/nnetar.html>)

En donde y_i representa el valor que ha de entrar a la función de activación; w_i es el peso estimado de los datos de entrada; x_i es el Input y b_i es el sesgo asociado a cada Input.

Posteriormente se tienen capas ocultas en las cuales se podrán modificar los pesos, esto por medio de una función de activación¹⁶ que recibe la combinación lineal previamente descrita y transforma los datos en una escala que usualmente va de (-1,1) o bien de (0,1), dicha función puede definirse como una función de salida $f(y_i)$ que transforma el estado actual de activación y_i a una señal de salida $s(y_i)$ (Villanueva Espinoza, 2002), hasta llegar a la capa de salida donde se tendrá el resultado esperado (Vidal Gonzalez, 2015). Existen diferentes funciones de activación que pueden ser utilizadas acorde al tipo de datos que se tenga disponible, algunas de las funciones de activación más comunes son las siguientes:

- Rectified Linear Unit (RELU): Como lo menciona Calvo en el artículo “Crear red Neuronal desde las matemáticas”, actualmente es una de las funciones de activación más utilizadas en el mundo para las capas ocultas dentro de la red y esta ha de activar la misma si y solo si el estado actual de activación es mayor a 0, la función se define como sigue:

$$relu(y) = \max(0, y)$$

- Función Sigmoide: Se ocupa por lo regular como una neurona de salida, sin embargo, no es limitante para ser utilizada como función dentro de las neuronas de la capa oculta, esta función evita la obtención de valores extremos ya que los resultados de salida existen entre 0 y 1.

$$sigmoide(y) = \frac{1}{1 + e^{-y}}$$

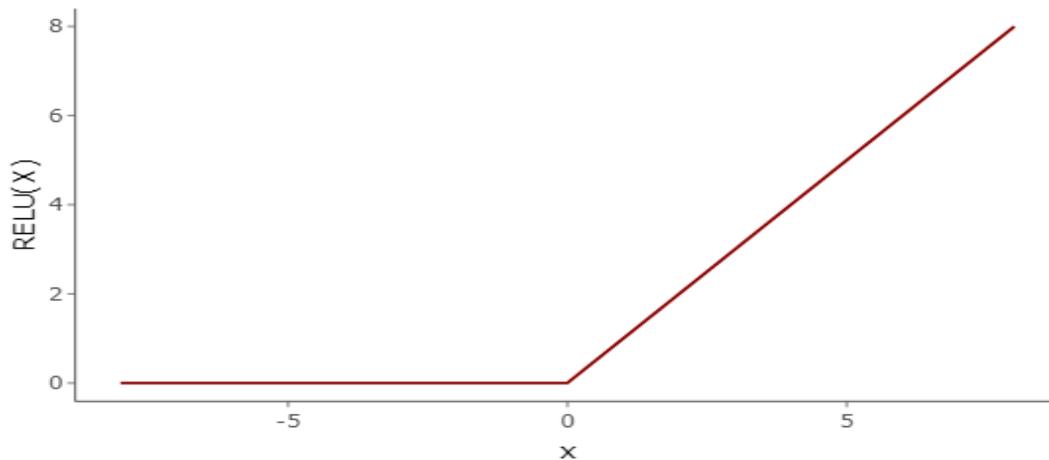
- Tangente hiperbólica: Esta función tiene un comportamiento muy similar a la función sigmoide, sin embargo, la diferencia es que los resultados de salida de dicha función están contenidos en un intervalo que va desde -1 a 1.

¹⁶ Como cada neurona tiene una función lineal se tiene el problema de que al final la suma de todo seguiría siendo una función lineal por lo que la función de activación transforma las salidas a fin de que no sean lineales.

$$\text{Tanh}(y) = \frac{1 - e^{-2y}}{1 + e^{-2y}}$$

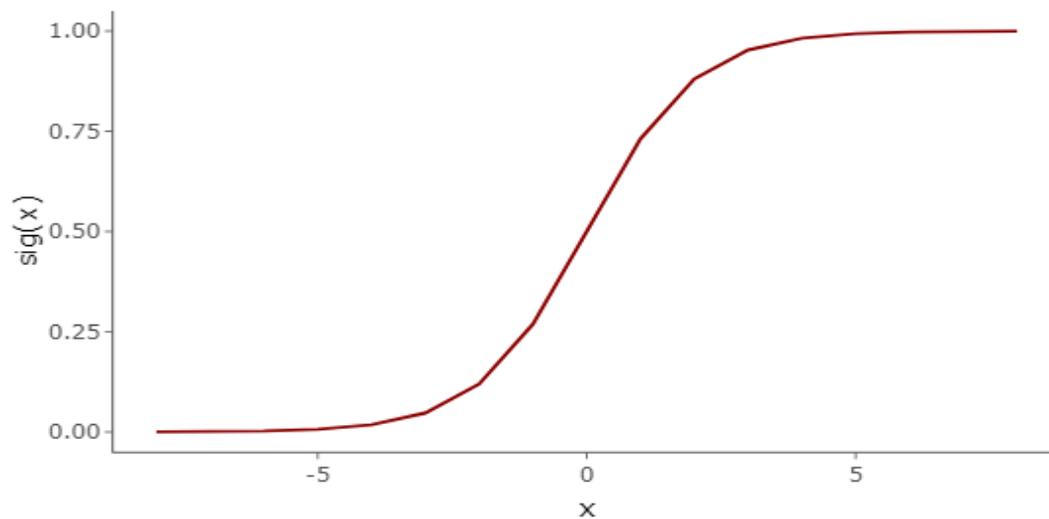
Una vez explicado el funcionamiento, así como la definición de cada una de las funciones de activación, a continuación, se muestra en los gráficos 3.2, 3.3 y 3.4 el comportamiento que cada una de estas funciones tiene.

Gráfico 3.2. Función de activación RELU.



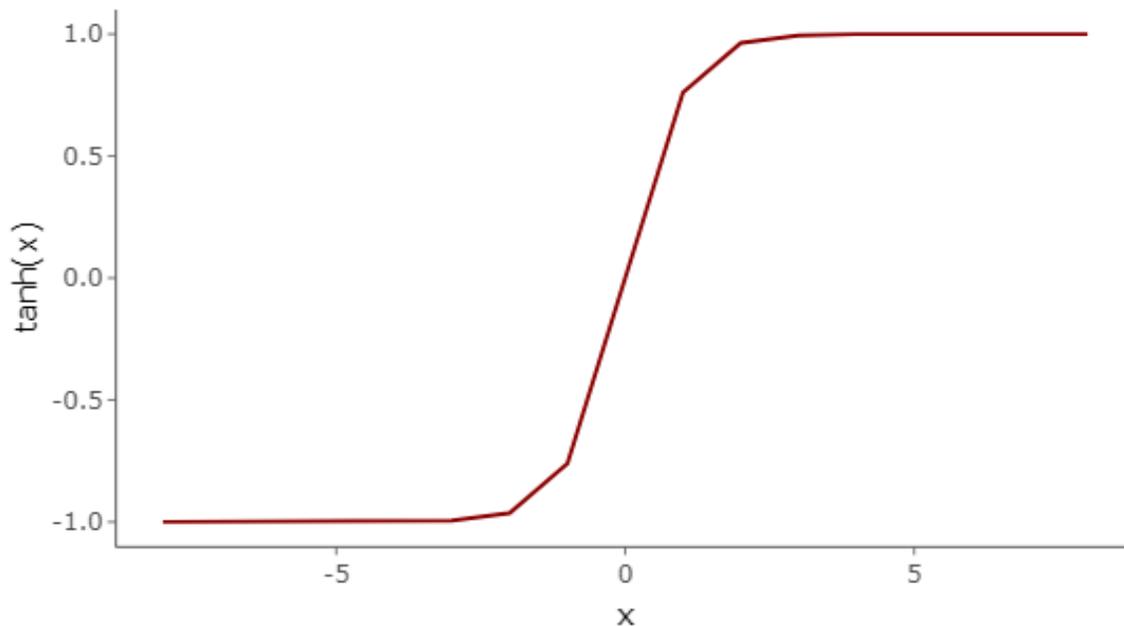
Fuente: Elaboración propia.

Gráfico 3.3. Función de activación Sigmoide.



Fuente: Elaboración propia.

Gráfico 3.4. Función de activación Tangente hiperbólica.



Fuente: Elaboración propia.

Durante el desarrollo, entreno y pronóstico de los resultados de una red neuronal interviene un concepto demasiado importante dentro del terreno de redes neuronales, este concepto es la función de coste, la cual permite cuantificar los errores de predicción cometidos por la red en cuestión, usualmente emplea la función del error cuadrado medio (MSE en inglés) así como la función del error absoluto medio (MAE en inglés). Lo que se busca al desarrollar una red neuronal es encontrar aquella red que minimice las medidas de error previamente descritas. Para fines de esta tesis emplearemos estas funciones de coste para encontrar aquella red que pronostique de mejor manera los datos del índice Nacional de Precios al Consumidor en México para posteriormente proyectar la inflación mensual anualizada en nuestro país.

Es importante explicar que el método de aprendizaje o principio que se ocupa para la parte de los perceptrones multicapa es a través del algoritmo de backpropagation cuyo fin es “conseguir propagar los errores cometidos por las neuronas de salida hacia atrás” cuya teoría fue aportada por Hinton y Williams del grupo PDP en 1986, cuya aportación indica que, conociendo el error pueden modificar los pesos las conexiones en dirección opuesta al gradiente para decrecer el incremento del error (Vidal Gonzalez, 2015).

En el ámbito de las series de tiempo existe una aplicación inmensa y popular de redes neuronales, sin embargo, una de las principales es aquella red que utiliza rezagos de la variable dependiente como variables que serán utilizadas en la capa de entrada como Inputs, a este tipo de redes se les conoce como Redes Neuronales Autorregresivas¹⁷ (Neural Network Autoregression) y estas se explican a continuación.

3.3.4.1 Red Neuronal Autorregresiva

Como ya se mencionó con anterioridad, las redes neuronales autorregresivas son un tipo de red que toma como Inputs aquellos valores rezagados de la serie de tiempo, el supuesto de este tipo de redes es que solamente existe una capa de entrada con n número de Inputs, una capa oculta y un solo Output, un ejemplo claro de estas puede ser el Gráfico 3.1 en donde se pueden observar 4 neuronas que conforman la capa de entrada, 3 neuronas que conforman la capa oculta y 1 un solo output. La esencia principal de este tipo de redes radica en tratar de explicar la variable dependiente en el tiempo t (y_t) en función de sus p rezagos ($y_{t-1}, y_{t-2}, \dots, y_{t-p}$), en adición se debe de tomar en cuenta el número de nodos o neuronas a utilizar dentro de la capa oculta, a esto se le conoce como k que a un inicio puede ser arbitraria, la notación sugerente para este tipo de red es la siguiente:

$$NNAR(p, k)^{18}$$

En donde, al tomar los p últimos rezagos de la serie de tiempo y establecemos un número k de neuronas que irán dentro de la capa oculta. Como dato importante a resaltar vemos que por el tipo de red que usa los rezagos de la variable de interés para explicar su propio comportamiento tendremos cierta similitud con los modelos autorregresivos (AR) para series de tiempo, siendo entonces que si solo se considera un rezago y 0 neuronas dentro

¹⁷ Este tipo de Red de igual modo sigue la arquitectura de una Red Neuronal multicapa alimentada hacia adelante.

¹⁸ Véase el capítulo 12, sección 12.4 del libro Forecasting: Principles and Practice de Rob. Hyndman donde se habla de este tipo de red. (<https://otexts.com/fpp3/nnetar.html>)

de la capa oculta (NNAR(1,0)), tendremos lo mismo que un modelo autorregresivo de orden 1 (AR(1)) sin tomar en cuenta el supuesto de estacionariedad en media.

Si los datos que se desea pronosticar tienen cierto componente estacional de igual modo puede considerarse como método de pronóstico el uso de una red neuronal autorregresiva considerando para estos casos la implementación de P rezagos estacionales de la serie, esto se traduce en tratar de explicar el comportamiento de la variable dependiente en el tiempo t (y_t) en función de sus p rezagos ($y_{t-1}, y_{t-2}, \dots, y_{t-p}$) y sus P rezagos estacionales ($y_{t-m}, y_{t-2m}, \dots, y_{t-Pm}$), de igual modo se debe considerar el definir un número k de neuronas dentro de la capa oculta, para este tipo de red, puede definirse un valor k como sigue:

$$k = \frac{p + P + 1}{2} \quad (1)$$

Estableciendo como k al entero más cercano resultado de la operación descrita en (1), la notación entonces para este tipo de red es la siguiente:

$$NNAR(p, P, k)_m^{19}$$

En donde los valores p y P representan la cantidad de rezagos y rezagos estacionales que servirán como Inputs para la capa de entrada, m usualmente representa el periodo estacional que sigue la serie y un número k de neuronas que integran la capa oculta de la red.

El proceso de pronóstico una vez definida la red que se ha de utilizar es sencilla, por ejemplo, en este caso en particular si queremos pronosticar el valor del INPC en enero de 2022, utilizando como datos de entrada los últimos $p = 2$ valores de este índice y $P = 1$ valores estacionales del mismo, los valores para la capa de entrada serían los siguientes:

$$INPC_{t-1}, INPC_{t-2}, INPC_{t-12}$$

¹⁹ Véase el capítulo 12, sección 12.4 del libro Forecasting: Principles and Practice de Rob. Hyndman donde se habla de este tipo de red.

En donde: $INPC_{t-1}$: Representa el valor del INPC en diciembre de 2021; $INPC_{t-2}$: Es el valor del INPC en noviembre de 2021; $INPC_{t-12}$: Se traduce a considerar el valor del INPC en enero de 2021

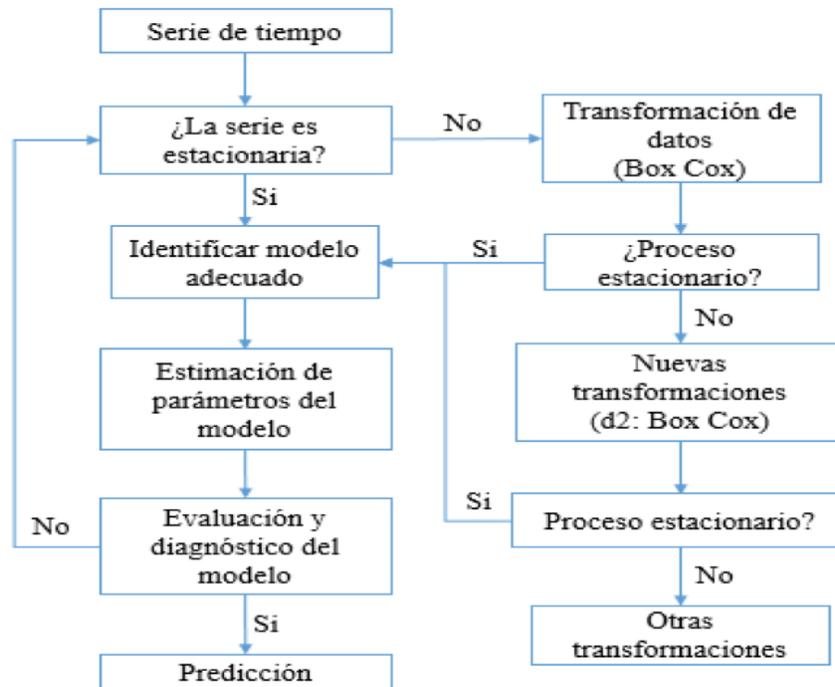
Así mismo tomando en cuenta la expresión descrita en (1) el valor óptimo de k sería igual a 2, que bien entonces sería el número de neuronas a considerar para este ejemplo dentro de la capa oculta. El proceso de pronóstico es iterativo, por lo que si queremos horizontes de pronóstico más lejanos se deben de tomar en cuenta los valores pronosticados, es decir, si se desea obtener pronósticos para febrero de 2022 utilizando este tipo de red, deberemos tomar en cuenta el valor de enero pronosticado, el proceso continúa hasta concluir con el horizonte temporal de pronóstico.

3.4 Comparación de modelos: SARIMA-GARCH; SARIMAX y Redes Neuronales

Una vez definidos los modelos que en el siguiente capítulo se han de desarrollar para llevar a cabo los pronósticos del INPC para así proyectar la inflación mensual anualizada en México, es conveniente presentar al lector el modo de desarrollo de cada modelo, así como también la forma en que este trabajo selecciona aquel que cometa menos errores de pronóstico.

Para la selección de la ecuación más representativa para pronosticar los datos, es común elegir entre diversas metodologías, sin embargo, para fines de esta investigación la metodología a utilizar será la desarrollada por Box y Jenkins, la cual consiste en una serie de pasos para elegir al modelo que comete menos errores a la hora de realizar un pronóstico para la serie de datos, esta metodología comienza preguntándose si la serie es estacionaria, en caso de no serlo es necesario aplicar una transformación (Box-Cox) acorde a los datos que se tienen así mismo una diferencia para controlar el problema de la no-estacionariedad, posteriormente es necesario identificar el modelo para luego encontrar los parámetros que mejor describan el comportamiento de nuestro conjunto de datos, luego como paso final se realizan los diagnósticos estadísticos pertinentes para elegir una ecuación candidata que realice los pronósticos, dichos pasos se resumen de mejor manera en el siguiente Gráfico.

Gráfico 3.5. Desarrollo de la metodología Box & Jenkins para la elección de un modelo de serie de tiempo.



Fuente: Fajardo Ortiz (2017).

Así mismo se consideran 4 criterios adicionales respecto a la elección del mejor candidato para el pronóstico, estos criterios se enlistan a continuación:

1. Criterio de Akaike (Tsay, s/a): Medida de bondad de ajuste la cual describe la relación existente entre el sesgo y la varianza de un modelo que en pocas palabras representa la exactitud y complejidad de este, el criterio se define como sigue:

$$AIC = \frac{2k}{t} - \frac{2}{t} \ln L$$

Donde k : Representa el número de parámetros del modelo; t : es el tamaño de la muestra; $\ln L$: es la función de log-verosimilitud para el modelo estadístico.

2. Error absoluto medio (Tsay, 2005): Utilizado en esta tesis para medir el promedio de los errores en valor absoluto, para su cálculo es necesario considerar lo siguiente:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i|$$

En donde, y_i : representa el valor que toma la serie (lo que se observa); \tilde{y}_i : el valor pronosticado para la serie. Y n : el número total de observaciones.

3. Raíz del error cuadrático medio (Tsay, 2005): La cual puede definirse como la desviación estándar de los residuales o errores de predicción, indica el nivel de concentración de los datos en la línea de mejor ajuste. Esta medida de precisión se obtiene como sigue:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2}$$

En donde se destaca la intervención de los valores observados y pronosticados de un modelo en particular.

4. Error porcentual absoluto: Permite de igual modo establecer un desempeño de pronóstico y mide el tamaño del error absoluto en términos porcentuales, para su cálculo se considera lo siguiente:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \tilde{y}_i|}{y_i}$$

Explicados los criterios de selección y desarrollo de modelos, en el capítulo siguiente se muestra la implementación de toda la metodología anteriormente descrita para el pronóstico del INPC en México para todo el año 2022.

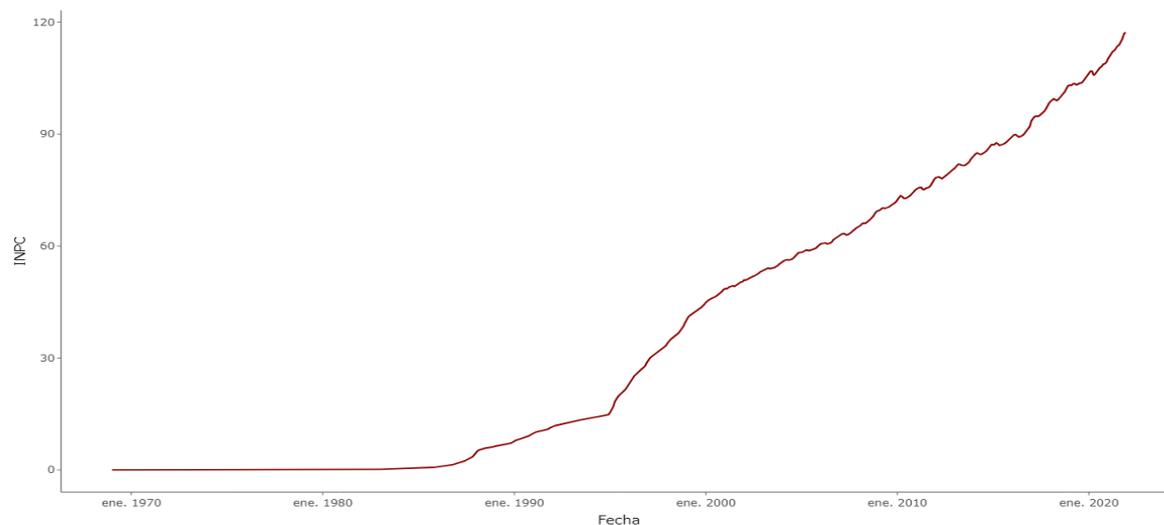
Capítulo 4. Resultados de la investigación

En este capítulo se presentan los resultados de los análisis desarrollados con los cuales se obtienen las comparaciones y los pronósticos del Índice Nacional de Precios al Consumidor. El orden de exposición de los resultados es: modelo SARIMA, SARIMAX, Red Neuronal y Red Neuronal Autorregresiva.

4.1 Modelo SARIMA

Como es común dentro de la metodología tradicional de series de tiempo debe tenerse en cuenta el comportamiento de la serie, para ello se genera el gráfico de acuerdo con los datos obtenidos de la página del INEGI. Cabe resaltar que la frecuencia de estos es mensual y comienza en enero de 1969 terminando en diciembre de 2021 teniendo un total de 636 registros para llevar a cabo la presente investigación.

Gráfico 4.1. Índice Nacional de Precios al Consumidor: enero de 1969 a diciembre 2021.



Fuente: elaboración propia con datos del INEGI utilizando la paquetería ggplot2 en RStudio.

Como se aprecia visualmente en el gráfico, el INPC tiene un comportamiento moderadamente constante al menos hasta principios de 1987, a partir de dicha fecha

comenzamos a apreciar un aumento a través del tiempo, es decir, vemos que existe una marcada tendencia positiva. Esto se explica claramente pues se espera que el valor del dinero a través del tiempo incremente su valor. Vemos también que el comportamiento de los datos pareciera una línea recta, sin embargo, no debemos descartar del todo las variaciones que son visibles a lo largo de la fecha de estudio pues seguramente como toda serie económica estemos frente a un posible efecto de volatilidad.

Por otro lado, es importante destacar que el Gráfico 4.1 desde la perspectiva metodológica tradicional de series de tiempo (Box-Jenkins) no cumple con la propiedad con la que toda serie de tiempo debería cumplir antes de ajustarse algún modelo: Esta propiedad es la estacionariedad en media. Es por ello por lo que se realizaron pruebas estadísticas para el conjunto de datos del INPC, obteniendo un valor de ($q = 8.712$, $p = 0.010$) para la prueba KPSS y valores del estadístico y p-value de ($q = -1.481$, $p = 0.798$) para la prueba DICKEY-FULLER

Como era de esperarse, en una primera instancia los datos no resultaron ser estacionarios y se confirma con los valores obtenidos de ambas pruebas estadísticas mencionadas con anterioridad. Para una mejor interpretación de los resultados a continuación se presenta un pequeño algoritmo que resume la manera en que son utilizadas las hipótesis al momento de llegar a la conclusión sobre la estacionariedad en media y/o tendencia. De acuerdo con las pruebas estadísticas de Kwiatkowski–Phillips–Schmidt–Shin y de Dickey-Fuller, se tomaron las siguientes decisiones.

a) Prueba Kwiatkowski–Phillips–Schmidt–Shin (KPSS):

1.- Definir las pruebas de hipótesis de la siguiente manera:

$$H_0 = \textit{La serie de tiempo es estacionaria en tendencia}$$

2.- Definir un nivel de significancia (α), para fines de esta investigación se establece un $\alpha = 0.05$.

3.- Si el valor p obtenido resulta ser menor que el valor de significancia ($p < \alpha$) entonces se rechaza la hipótesis nula, concluyendo que la serie de tiempo no es estacionaria en tendencia. Por otro lado, si el valor p obtenido resulta ser mayor al nivel de significancia ($p > \alpha$) no se tendría evidencia suficiente para rechazar la hipótesis nula.

b) Prueba Dickey-Fuller:

1.- Definir las pruebas de hipótesis de la siguiente manera:

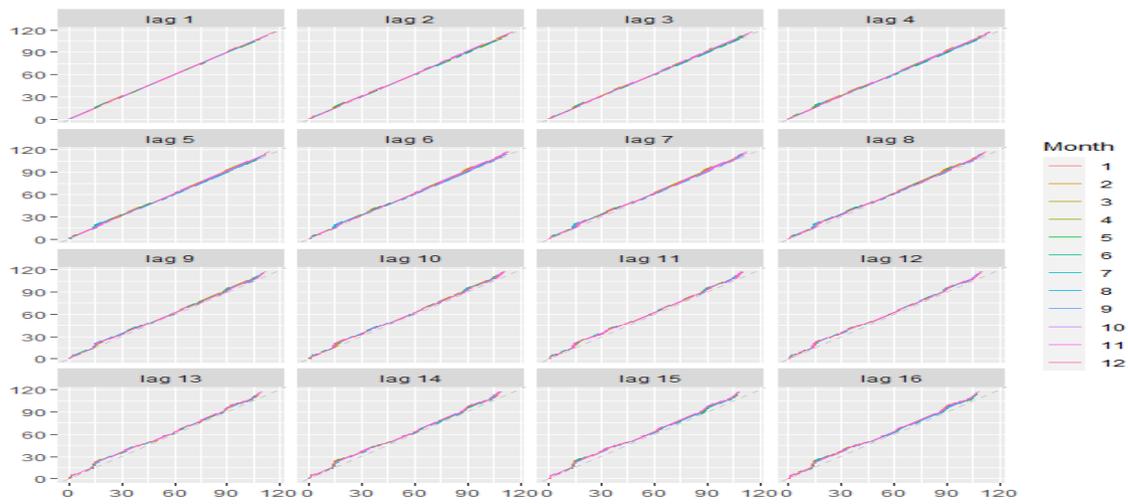
$$H_0 = \textit{La serie de tiempo no es estacionaria}$$

2.- Definir un nivel de significancia (α), para fines de esta investigación se establece un $\alpha = 0.05$

3.- Si el valor p obtenido resulta ser menor que el valor de significancia ($p < \alpha$) entonces se rechaza la hipótesis nula y se concluye que la serie de tiempo es estacionaria, es decir, no presenta un cierto grado de dependencia, así como también su varianza podría decirse constante a través del tiempo. Por otro lado, si el valor p obtenido resulta ser mayor al nivel de significancia ($p > \alpha$) no se tendría evidencia suficiente para rechazar la hipótesis nula y por lo tanto la serie no resulta estacionaria, presenta dependencia en sus valores, así como también el tema de la varianza que no resulta ser constante.

Con el algoritmo recientemente descrito puede afirmarse que la serie del INPC presenta un cierto grado de dependencia en sus valores que se manifiesta en forma de tendencia, así como también que su varianza no es constante a lo largo del tiempo, por ende, la serie no es estacionaria, el resultado de las pruebas anteriormente realizadas puede corroborarse en la tabla A.1 en el apartado de anexos. Para confirmar dicha tendencia se presenta el Gráfico 4.2 que representa una comparativa de los valores del INPC vs los primeros 16 rezagos de la misma variable.

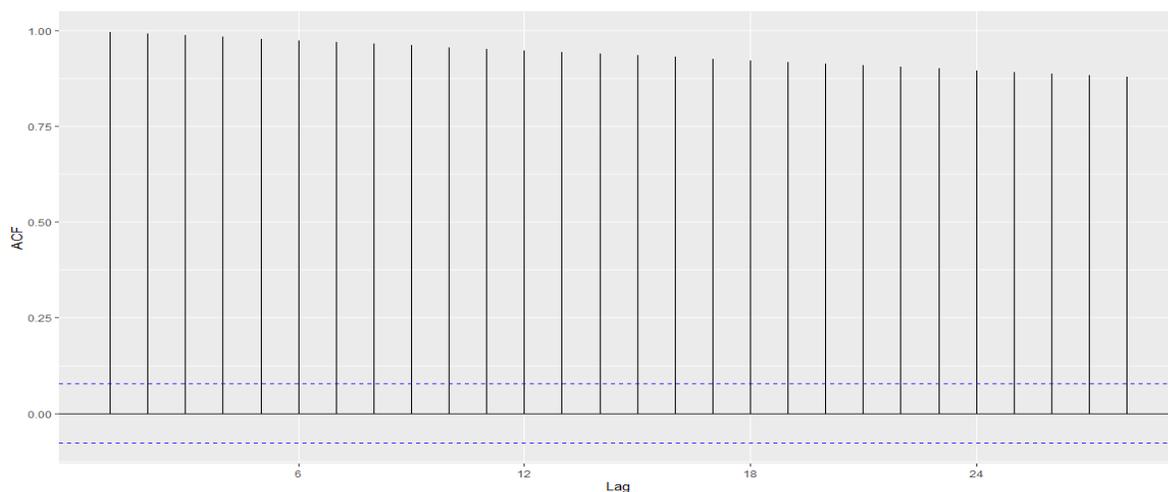
Gráfico 4.2. Valores del INPC vs sus primeros 16 rezagos.



Fuente: Elaboración propia con datos del INEGI y el comando gglagplot() de RStudio.

Una vez confirmada la falta de estacionariedad resulta buena práctica corregir el efecto de la tendencia, así como también la poca reversión a la media de la serie. Para ello se procede a realizar las primeras diferencias de los datos para que logren alcanzar las propiedades estadísticas necesarias antes de ajustarles cualquier modelo. La necesidad de solo considerar primeras diferencias se ve reflejada en el Gráfico 4.3 en donde se ve la persistencia en los rezagos conforme el número de estos incrementa, teniendo así un decaimiento bastante lento hacia cero.

Gráfico 4.3. Función de Autocorrelación del Índice Nacional de Precios al Consumidor



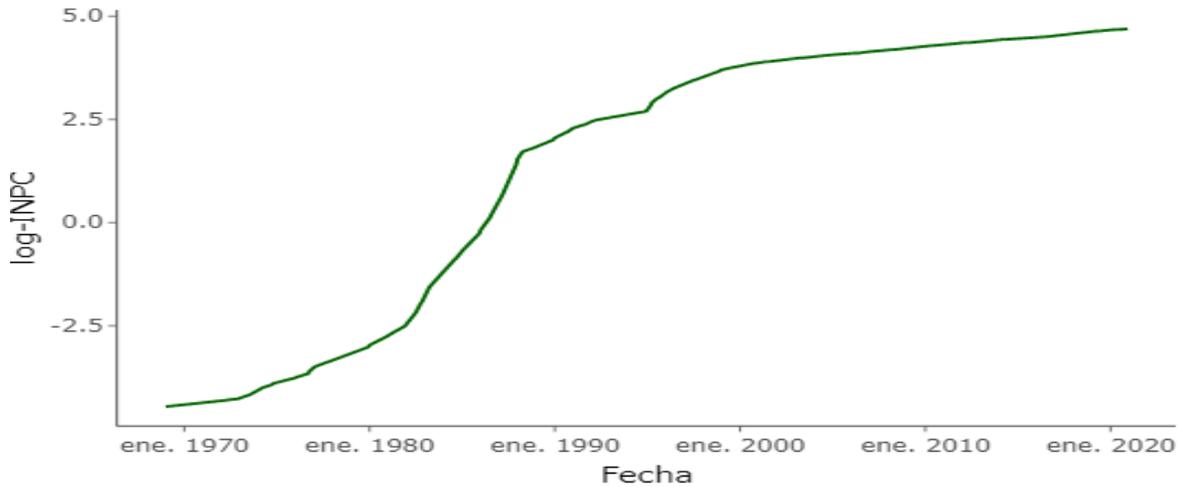
Fuente: Elaboración propia.

Sin embargo, antes de continuar con el tratamiento de la serie es importante aclarar los métodos que se han de utilizar a partir de este momento para poder ajustar un modelo de serie de tiempo, en primer instancia y respaldados por la propuesta del profesor Hyndman (Hyndman & Athanasopoulos, *Forecasting: Principles and Practice*, 2022) se han de separar el conjunto de datos del INPC en dos partes: la primera será denominada como nuestro conjunto de entrenamiento, el cual servirá para ajustar cualquier parámetro y/o ecuación para realizar los pronósticos. La siguiente parte será utilizada para probar la precisión de cada modelo que el conjunto de entrenamiento pueda crear, esta parte de los datos se denominará conjunto de prueba.

Usualmente para definir el tamaño de cada conjunto se utiliza el 20% del total de la muestra para el conjunto de prueba y el 80% restante para el conjunto de entrenamiento. Sin embargo, también es posible definir el tamaño del conjunto de entrenamiento con base en el horizonte temporal de pronóstico que se desea realizar, para fines de este trabajo se toman los datos de enero de 1969 hasta diciembre de 2020 como conjunto para entrenar diversos modelos y a partir de enero de 2021 hasta diciembre de ese mismo año como aquel conjunto de prueba que será de utilidad para probar la precisión de los modelos.

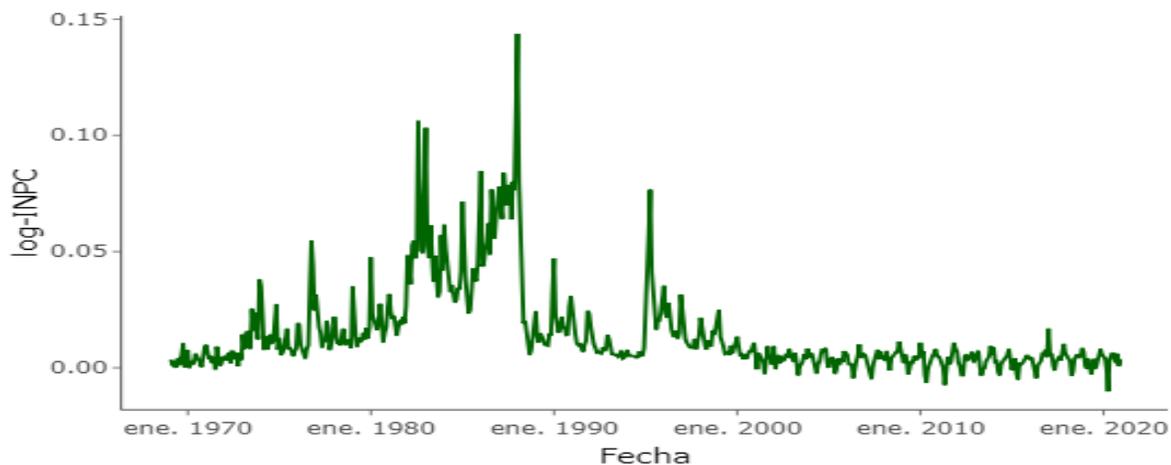
En adición también se realiza una transformación logarítmica respaldados por el artículo publicado por Box y Cox (Pengfei, 2005) lo anterior para tener nuestros datos dentro de una misma escala, así como también para temas relacionados con la varianza de la serie, ergo, se espera que con este ajuste nuestra serie tenga un comportamiento más estable: en términos sencillos lo que se busca con esta transformación es normalizar los datos. El Gráfico 4.4 muestra la transformación aplicada a nuestros datos de entrenamiento y en adición se muestra en el Gráfico 4.5 las primeras diferencias de la serie de tiempo transformada.

Gráfico 4.4. Transformación Logarítmica del INPC en México (Conjunto de entrenamiento).



Fuente: Elaboración propia.

Gráfico 4.5. Primeras diferencias del conjunto de entrenamiento.



Fuente: Elaboración propia.

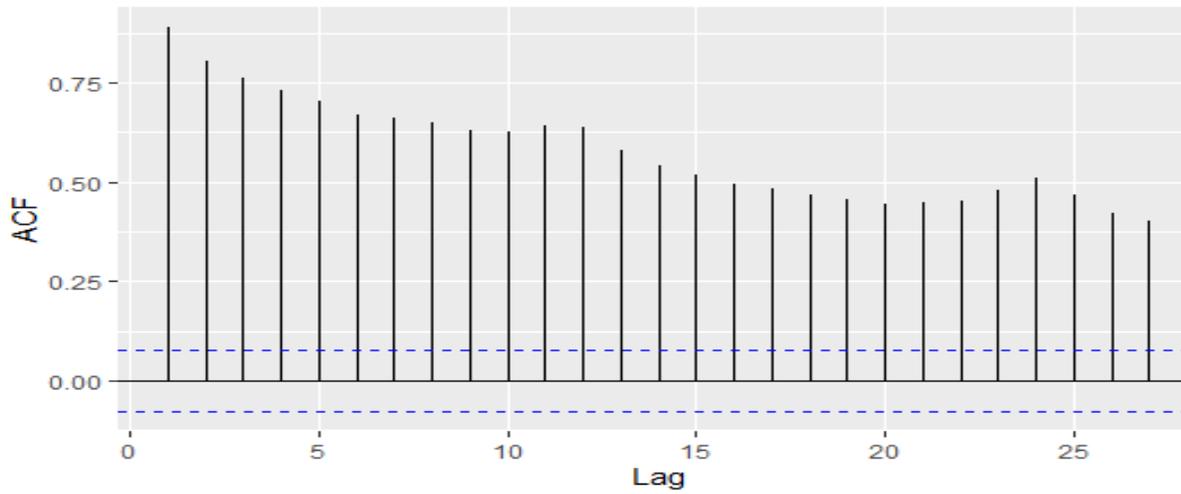
En el gráfico 4.4 se observa que al aplicar la transformación el objetivo de estabilizar los datos se cumple. El tema de la varianza se mantiene más controlado, pues los valores ahora oscilan en una escala menor a diferencia del Grafico 4.1. Se destaca también que las primeras diferencias del conjunto de entrenamiento comienzan a mostrar esa reversión a la media que pudiese ser sinónimo de estacionariedad, es importante mencionar que el

Gráfico 4.5 muestra algunas propiedades de suma importancia y estas se resumen en los siguientes puntos:

1. Se aprecian oscilaciones a lo largo de la temporalidad del conjunto de entrenamiento, esto puede tener cierta relación con lo mencionado por Heath (Inflación, 2012) en donde se encuentra respaldo para deducir un cierto tipo de estacionalidad en la serie.
2. Existe un comportamiento de oscilaciones mayores en los periodos comprendidos de 1982 a 1988 sexenio en el cual el PIB per cápita tuvo una reducción anual de 2.1% en términos reales, además de un aumento del 90% en los precios al consumidor en términos anuales y el tipo de cambio tuvo incrementos mayores al 100% (Economía UNAM, 2010). Posterior a este sexenio los rezagos de la crisis que arrastraba México quedaron marcados al menos hasta el año 2000, pues el ritmo de crecimiento económico no fue igual desde entonces. Motivados por estos sucesos y debido a las características de las series económicas se opta por considerar a los modelos GARCH por el tema de la volatilidad.
3. En los inicios de los años 2000 existe una mayor estabilidad (Heath, 2012), pues desde esa fecha se establecen medidas con mayor rigor para controlar los efectos inflacionarios por parte de BANXICO, estableciendo una tasa objetivo que comprende un intervalo del 3% +- 1%.

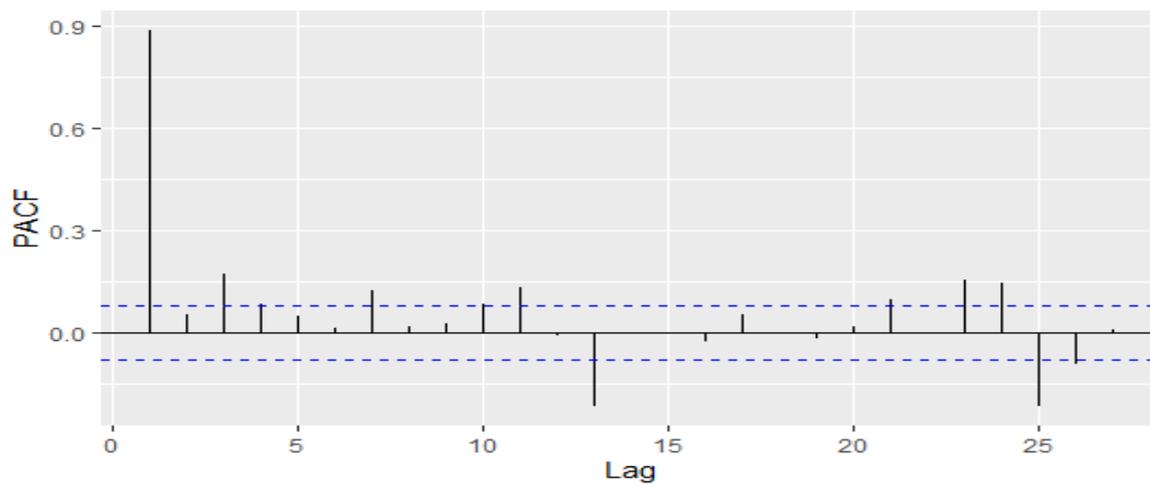
Es buena práctica seguir observando el gráfico de la función de autocorrelación (ACF en inglés) y la función de autocorrelación parcial (PACF en inglés) del conjunto de entrenamiento. Las gráficas 4.6 y 4.7 en primeras diferencias para dar mayor respaldo al punto 1, por última instancia se realiza también la descomposición de la serie (transformada, sin diferenciar) desde el enfoque multiplicativo para determinar si existe dicha estacionalidad (Gráfico 4.8).

Gráfico 4.6. ACF de las primeras diferencias del conjunto de entrenamiento con transformación logarítmica.



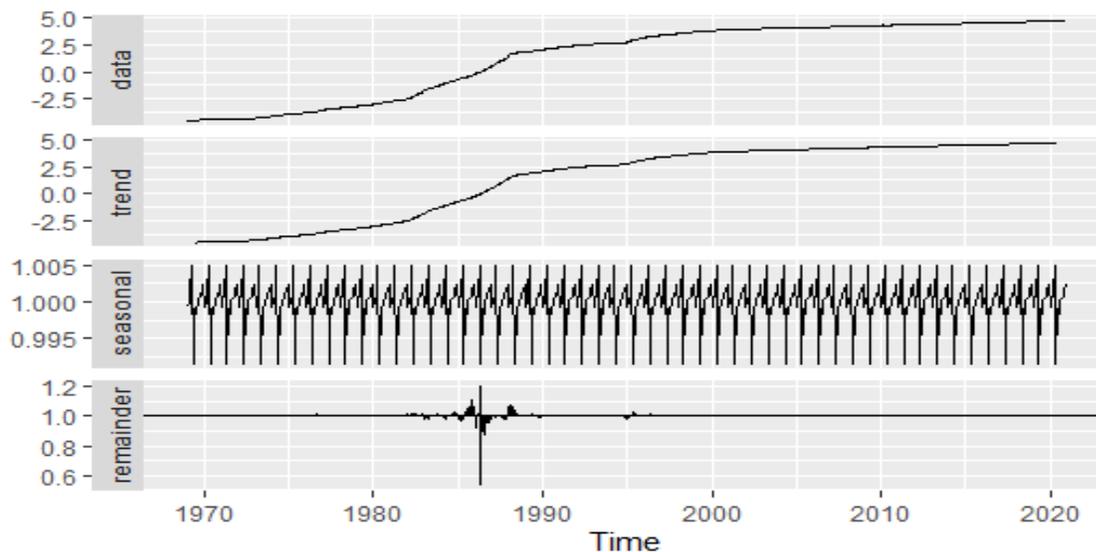
Fuente: Elaboración propia.

Gráfico 4.7. PACF de las primeras diferencias del conjunto de entrenamiento con transformación logarítmica.



Fuente: Elaboración propia.

Gráfico 4.8. Descomposición del conjunto de entrenamiento desde el enfoque Multiplicativo.

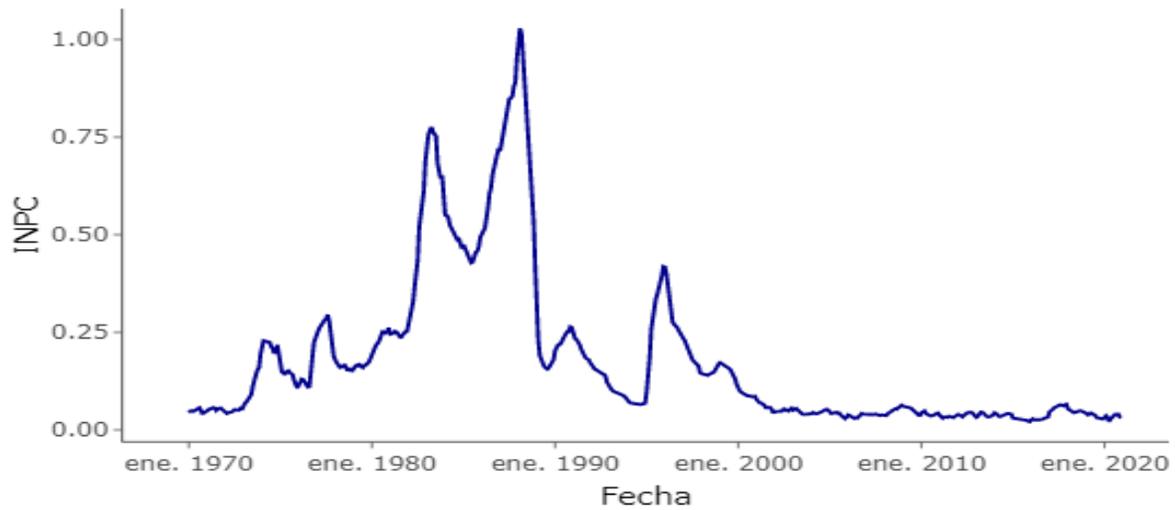


Fuente: Elaboración propia.

Como se puede apreciar, el tema estacional logra confirmarse pues, en el Gráfico 4.6 observamos como los rezagos de la serie decaen y justo incrementan su valor en los rezagos múltiplos de 12. En adición, en el Gráfico 4.8 se observa el efecto estacional que contiene el conjunto de entrenamiento, una vez teniendo evidencia suficiente que confirma la estacionalidad podemos incluir diferencias estacionales para suavizar nuestro conjunto de entrenamiento y volver estacionaria nuestra serie.

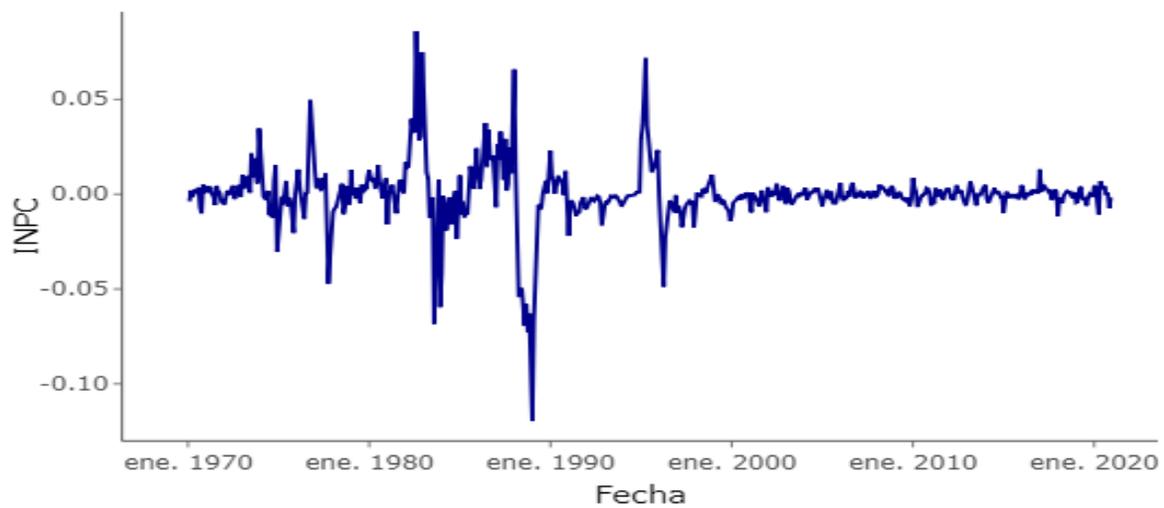
En conclusión, se debe diferenciar una vez para eliminar los efectos de la tendencia y se diferencia estacionalmente para suavizar los efectos de la estacionalidad del conjunto de entrenamiento. Para efectos de ilustrar la conclusión previa se presenta el gráfico 4.9 que muestra el efecto de diferenciar estacionalmente el conjunto de entrenamiento. Por otra parte, el gráfico 4.10 que permite observar el proceso diferenciado completamente (primeras diferencias y diferencias estacionales), mientras que el gráfico 4.11 que ilustra la función de autocorrelación que será base para comenzar con la definición de los modelos SARIMA. Por último, el gráfico 4.12 que se utilizó como apoyo para la definición de las ecuaciones a definir.

Gráfico 4.9. Diferencias estacionales aplicadas al conjunto de entrenamiento.



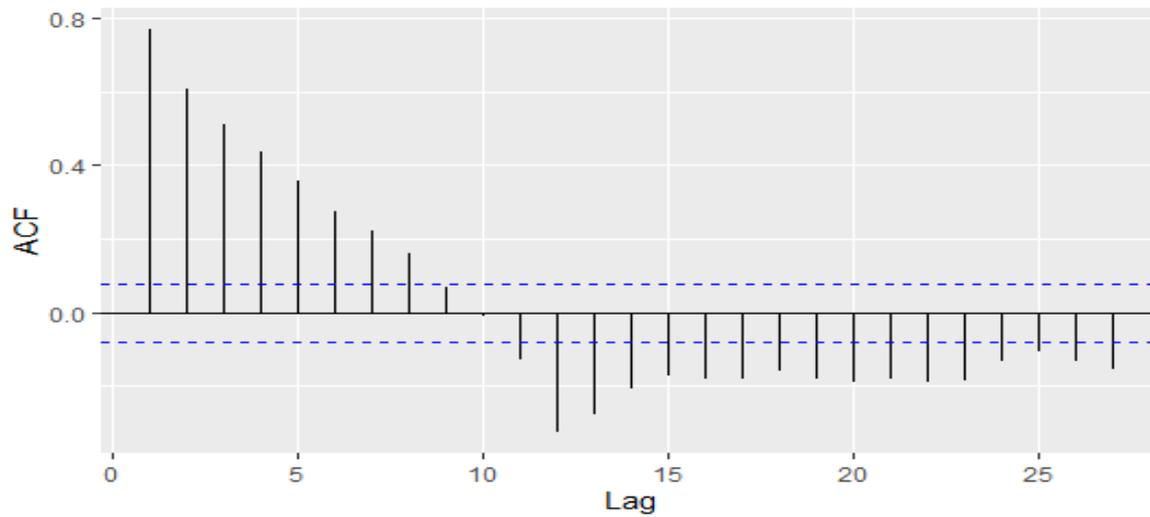
Fuente: Elaboración propia.

Gráfico 4.10. Proceso diferenciado completamente aplicado al conjunto de entrenamiento.



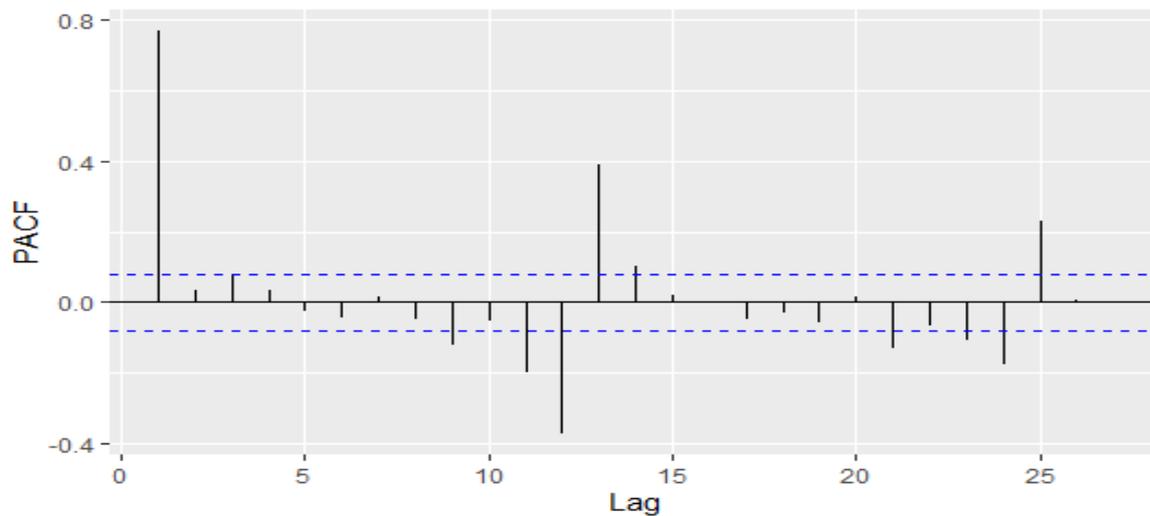
Fuente: Elaboración propia.

Gráfico 4.11. ACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento.



Fuente: Elaboración propia.

Gráfico 4.12. PACF del proceso diferenciado completamente aplicado al conjunto de entrenamiento.



Fuente: Elaboración propia.

Con este tratamiento a los datos de entrenamiento puede concluirse que la serie es estacionaria, esto debido a los resultados obtenidos por las pruebas KPSS y DICKEY-FULLER ($q = 0.0082$, $p = 0.10$; $q = -6.956$, $p = 0.010$) respectivamente. Los resultados

pueden verse en la tabla A.2 en la parte de anexos en donde se muestra un resumen de los estadísticos aplicados a los datos de entrenamiento diferenciados en primer orden y diferenciados estacionalmente.

Bajo el mismo algoritmo descrito en la interpretación de los resultados para la tabla A.1 puede decirse que existe evidencia suficiente para confirmar la estacionariedad en el conjunto de entrenamiento.

El paso siguiente es ajustar un modelo de serie de tiempo que contenga la propiedad de ser un modelo que respete el principio de parsimonia, es decir, un modelo sencillo que explique la mayor parte de los datos con el menor número de parámetros. Es importante mencionar algunas características que se consideran al momento de elegir a un modelo de serie de tiempo como el adecuado para fines de pronóstico en esta investigación, las características utilizadas se resumen a continuación en el siguiente listado:

1. Se busca que los coeficientes de la ecuación sean estadísticamente diferentes de 0, pues no tendría sentido incluir un coeficiente que no sea capaz de explicar el conjunto de datos.
2. Los residuos de la ecuación deben ser independientes y por ende no debe existir una correlación serial, para ello hemos prestado especial atención al valor p de la prueba Ljung-Box: $B = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n-k} \right) \sim \chi^2$ con *m* grados de libertad (Gujarati & Porter, 2010). Este valor deberá ser mayor a 0.05 para poder aceptar la hipótesis nula la cual nos dice que los residuos se distribuyen de manera independiente.
3. El criterio de Akaike (valor AIC) debe ser el mínimo entre ecuaciones, ya que con dicho criterio se expresa la bondad de ajuste del modelo resultante, así como su complejidad y manera de establecer una relación entre el sesgo y la varianza.
4. Se busca encontrar la medida de error más pequeña a la hora de medir la precisión de los pronósticos. Para ello, se centra el análisis en la medida del error absoluto medio (MAE en inglés), la raíz del error cuadrado medio (RMSE en inglés) y el error porcentual absoluto (MAPE en inglés) para elegir al mejor candidato para realizar los pronósticos a 1 año.

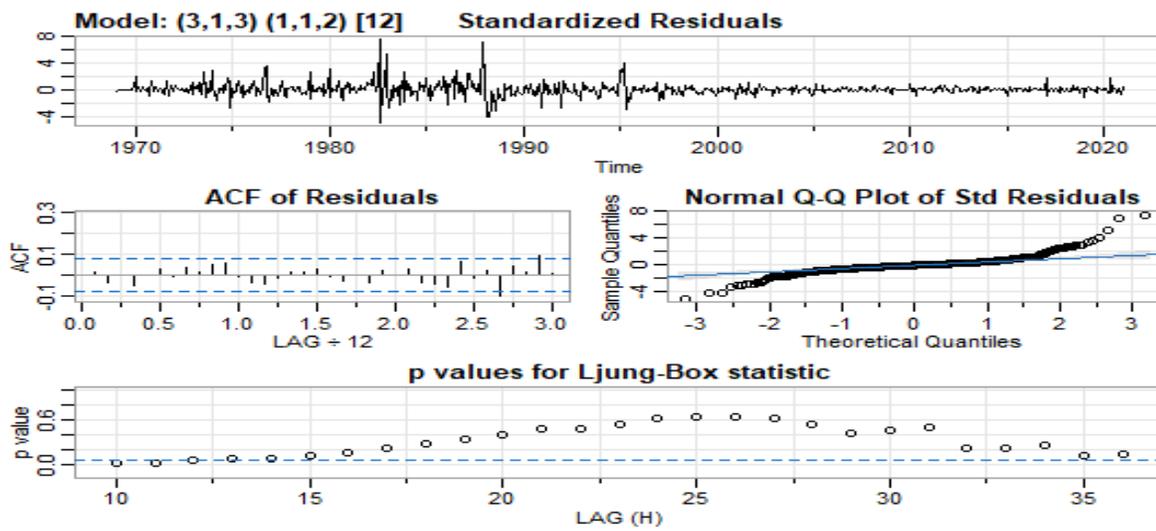
Debido al comportamiento de los rezagos en los Gráficos 4.11 y 4.12 se formula la ecuación $SARIMA(3,1,3)(1,1,2)_{12}$ la cual en primera instancia mostró los resultados

deseados a la hora de evaluar la calidad del modelo. Los coeficientes resultaron ser en su mayoría significativamente diferentes de cero excepto el coeficiente autorregresivo estacional (SAR), esto debido a que su estadístico t resultó ser -0.68 y su valor p cercano a 0.5. En adición, se observa que los residuales se comportan de una buena manera, pasando la prueba de hipótesis para la prueba Ljung-Box, obteniendo un valor p de 0.61. Respaldando este punto, los residuales están dentro de las bandas de confianza para la función de autocorrelación, así como también que los residuales de este primer modelo se comportan como si fuesen ruido blanco aproximándose a una distribución Normal. El Gráfico 4.13 muestra gráficamente el desempeño del modelo respecto a los residuos.

Resultado de la poca significancia del coeficiente autorregresivo para la parte estacional se decidió eliminarlo de la ecuación teniendo ahora un modelo $SARIMA(3,1,3)(0,1,2)_{12}$, con esta implementación vemos que todos los coeficientes resultan ser significativos, el valor p para la prueba Ljung-Box aumenta teniendo ahora un valor de 0.672, los residuos siguen teniendo ese comportamiento de ruido blanco y con aproximaciones a una distribución Normal. El Gráfico 4.14 demuestra el desempeño del modelo anteriormente descrito.

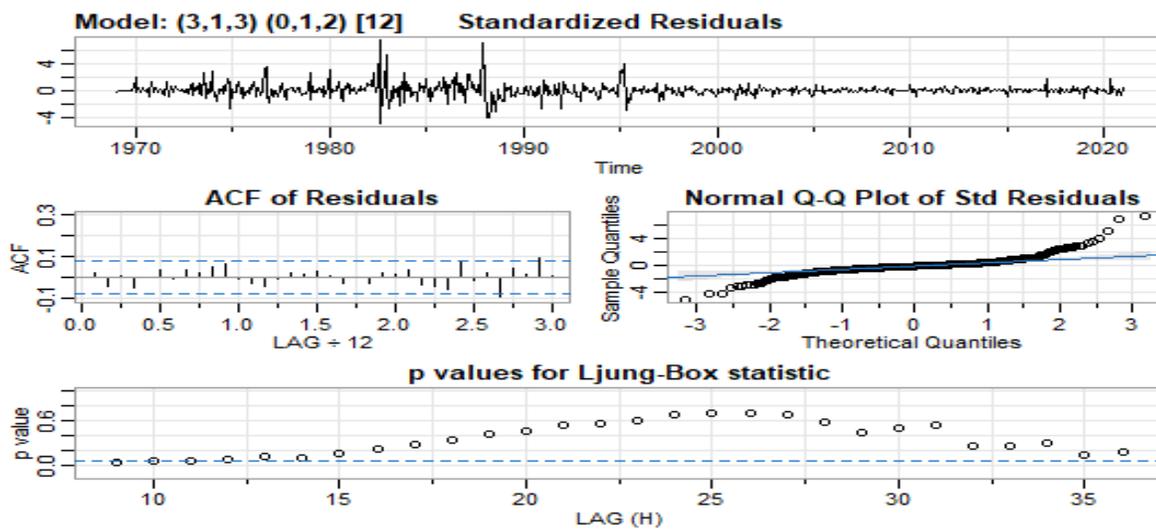
El paso siguiente fue encontrar una mejor ecuación en términos de parámetros ya que con las dos anteriormente descritas se tenía un total de $p + q + P + Q = 9$ para el modelo $SARIMA(3,1,3)(1,1,2)_{12}$ y un total de $p + q + P + Q = 8$ para el modelo $SARIMA(3,1,3)(0,1,2)_{12}$ por ende se decidió que una posible mejora a estos modelos era reducir en 1 unidad el termino de promedio móvil estacional dando origen a la ecuación $SARIMA(3,1,3)(0,1,1)_{12}$ (Gráfico 4.15) con un total de 7 parámetros mostrando de igual modo un buen desempeño al menos respecto a sus propiedades estadísticas.

Gráfico 4.13. Residuales del modelo $SARIMA(3, 1, 3)(1, 1, 2)_{12}$.



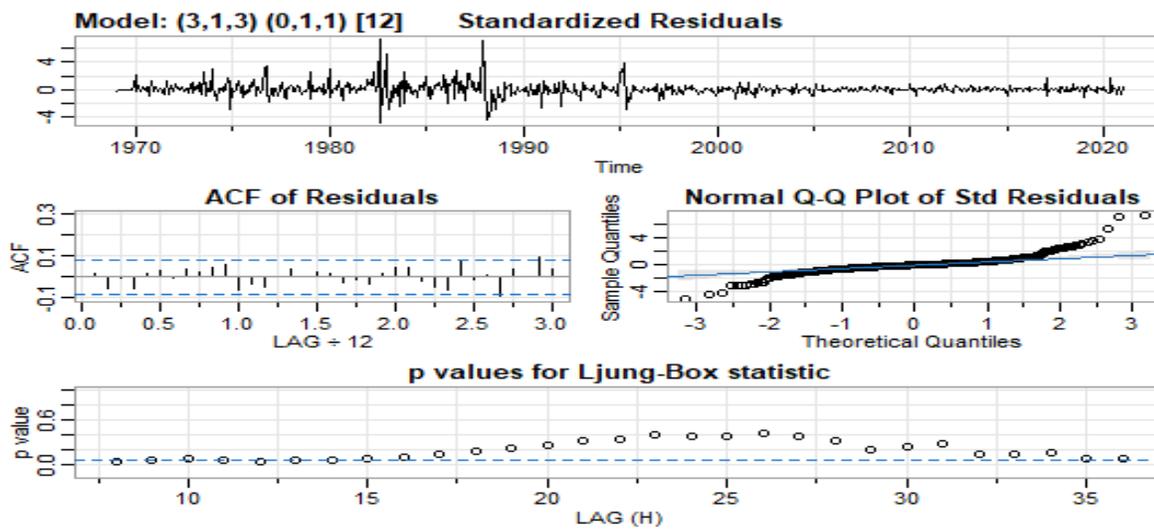
Fuente: Elaboración propia.

Gráfico 4.14. Residuales del modelo $SARIMA(3, 1, 3)(0, 1, 2)_{12}$.



Fuente: Elaboración propia.

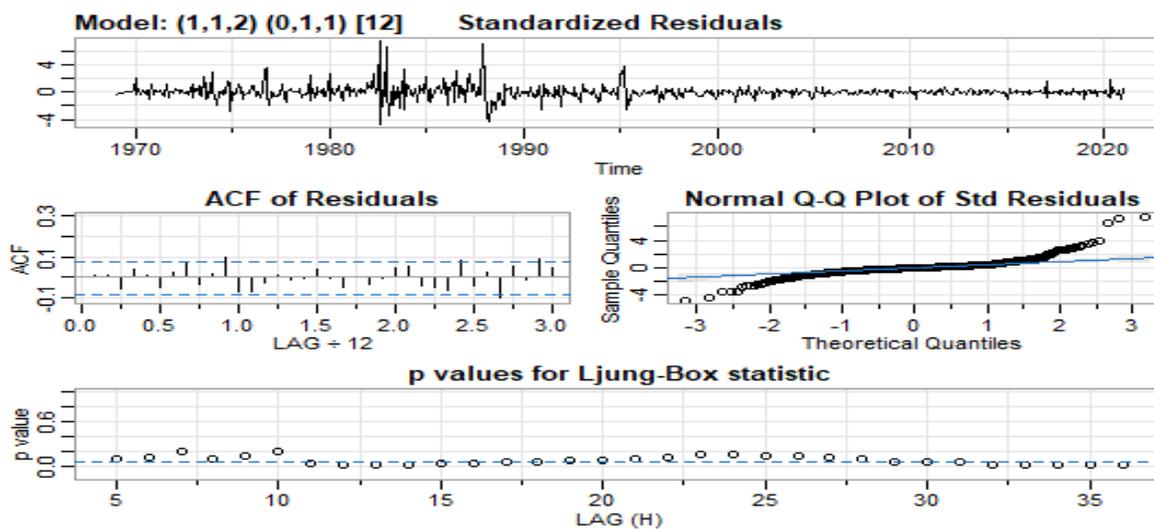
Gráfico 4.15. Residuales del modelo $SARIMA(3, 1, 3)(0, 1, 1)_{12}$.



Fuente: Elaboración propia.

De la misma manera, se optó por tener un modelo aún más sencillo, por ende aprovechando la tendencia del conjunto de entrenamiento y observando que la serie se explica en buena medida por sus errores pasados se optó por ajustar un modelo con 4 parámetros. La ecuación resultante es un modelo $SARIMA(1,1,2)(0,1,1)_{12}$ cuyas propiedades estadísticas siguen siendo óptimas. El Gráfico 4.16 muestra gráficamente el comportamiento de los residuales del modelo de 4 parámetros.

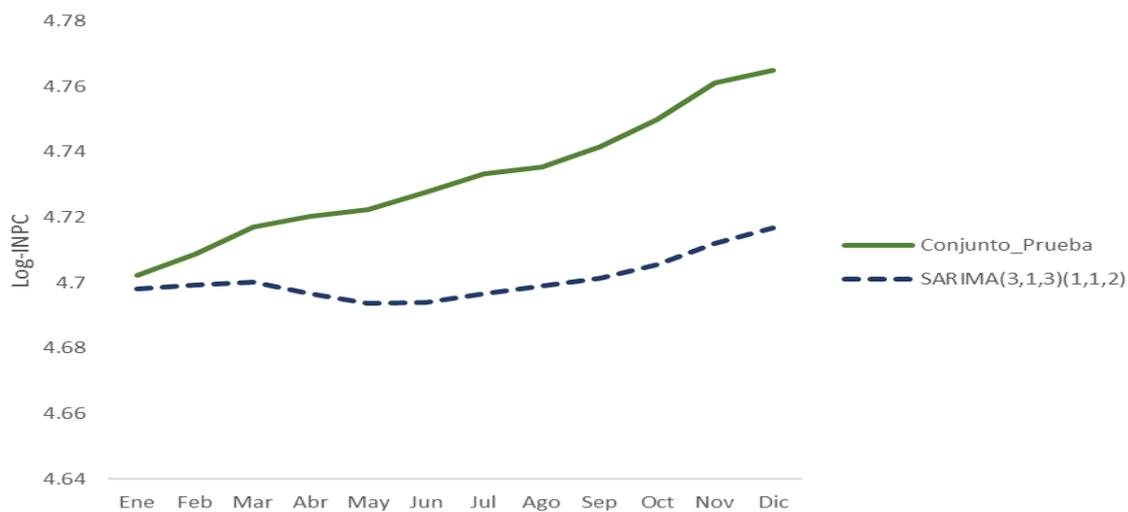
Gráfico 4.16. Residuales del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$.



Fuente: Elaboración propia.

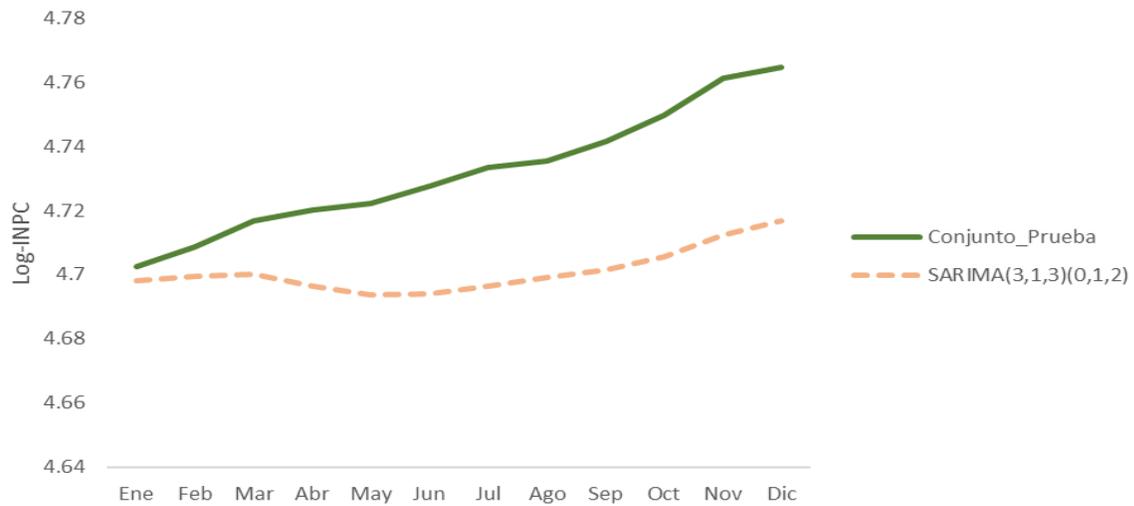
Una vez teniendo a los 4 candidatos el paso siguiente es probar la precisión de cada uno de estos modelos, para ello se pronostica el año 2021 (conjunto de prueba) y con base en los pronósticos se obtiene dicha precisión eligiendo para ello las medidas RMSE, MAE y MAPE más adecuadas (aquellas que minimicen el error). Para tener 1 modelo candidato en lugar de 4 y así poder realizar el pronóstico de todo el año 2022. A continuación, se muestra el conjunto de pronósticos que se han de comparar contra los valores reales del conjunto de prueba (Gráficos 4.17, 4.18, 4.19, 4.20) en el cual se define un candidato que destaca sobre los otros, este mismo resulta ser la ecuación $SARIMA(1,1,2)(0,1,1)_{12}$. Posteriormente, en la Tabla 4.3 se muestran las medidas de precisión RMSE, MAE y MAPE para cada candidato, así como su criterio de información (AIC).

Gráfico 4.17. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(1, 1, 2)_{12}$.



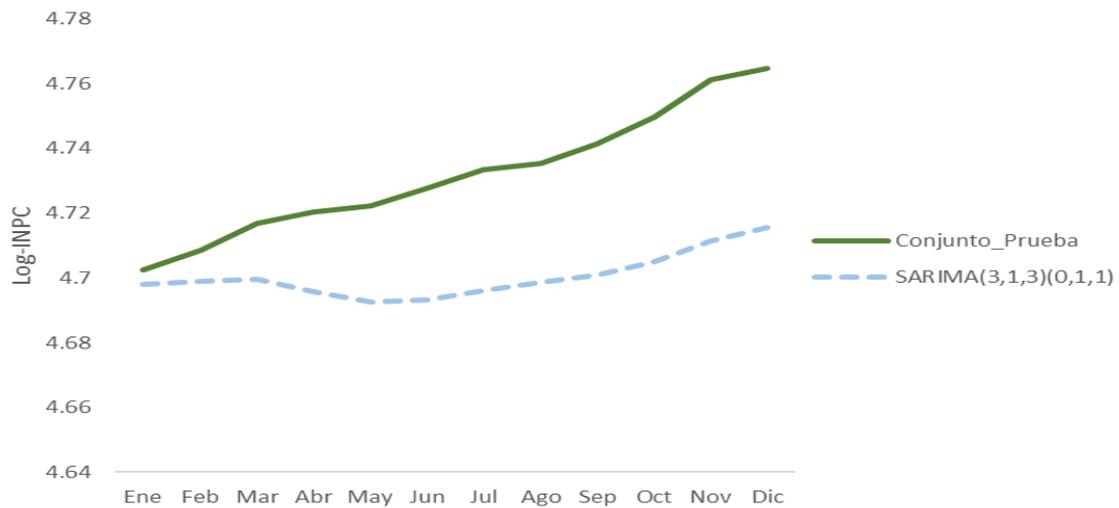
Fuente: Elaboración propia.

Gráfico 4.18. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(0, 1, 2)_{12}$.



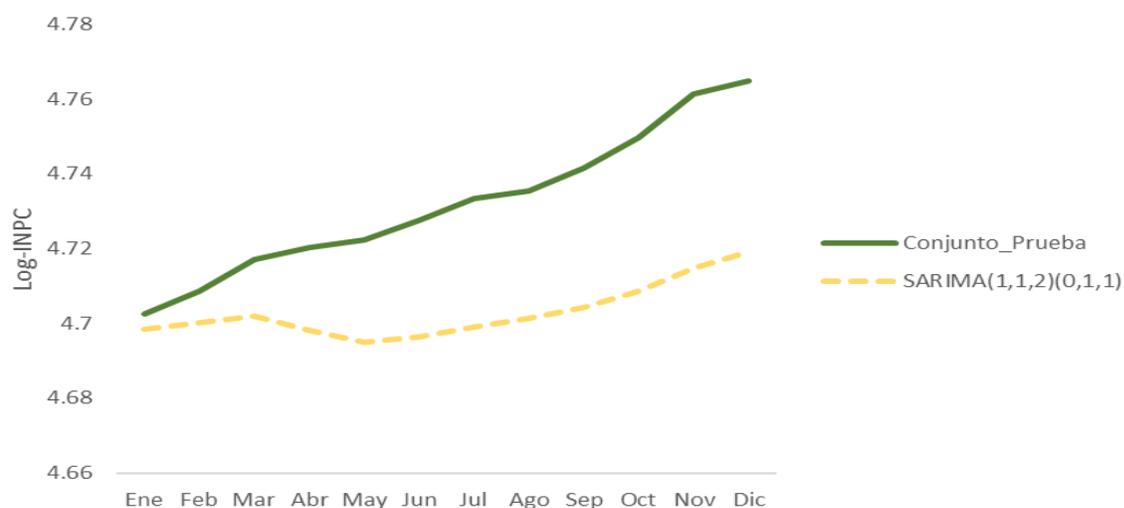
Fuente: Elaboración propia.

Gráfico 4.19. Comparativa: Conjunto de prueba vs modelo $SARIMA(3, 1, 3)(0, 1, 1)_{12}$.



Fuente: Elaboración propia.

Gráfico 4.20. Comparativa: Conjunto de prueba vs modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$.



Fuente: Elaboración propia.

Una vez conocidas las representaciones gráficas, en las Tablas 4.1, 4.2, 4.3, 4.4 y 4.5 muestran un resumen general de todas las características de los modelos explicados con anterioridad.

Tabla 4.1. Medidas de error para las 4 ecuaciones candidatas

Modelo	RMSE	MAE	MAPE	AIC
SARIMA(3,1,3)(1,1,2)	0.034	0.031	0.651	-4167.34
SARIMA(3,1,3)(0,1,2)	0.034	0.031	0.651	-4169.01
SARIMA(3,1,3)(0,1,1)	0.035	0.032	0.665	-4167.01
SARIMA(1,1,2)(0,1,1)	0.032	0.029	0.610	-4156.92

Fuente: Elaboración propia.

Tabla 4.2. Coeficientes y significancia del modelo $SARIMA(3, 1, 3)(1, 1, 2)_{12}$

Coeficiente	Valor	Estadístico t	p-value
AR1	0.483	4.209	0.000
AR2	-0.409	-5.284	0.000
AR3	0.761	13.274	0.000
MA1	0.2851	2.442	0.015
MA2	0.671	8.132	0.000
MA3	-0.258	-5.045	0.000
SAR1	-0.214	-0.684	0.494
SMA1	-0.653	-2.095	0.037
SMA2	-0.072	-0.281	0.779

Fuente: Elaboración propia.

Tabla 4.3. Coeficientes y significancia del modelo $SARIMA(3, 1, 3)(0, 1, 2)_{12}$

Coeficiente	Valor	Estadístico t	p-value
AR1	0.483	4.198	0.000
AR2	-0.399	-6.399	0.000
AR3	0.767	11.925	0.000
MA1	0.304	2.678	0.008
MA2	0.682	7.208	0.000
MA3	-0.259	-5.079	0.000
SMA1	-0.860	-19.651	0.000
SMA2	0.094	1.994	0.047

Fuente: Elaboración.

Tabla 4.4. Coeficientes y significancia de estos del modelo $SARIMA(3, 1, 3)(0, 1, 1)_{12}$

Coeficiente	Valor	Estadístico t	p-value
AR1	0.412	6.3041	0.000
AR2	-0.388	-7.925	0.000
AR3	0.790	14.095	0.000
MA1	0.356	4.569	0.000
MA2	0.718	8.567	0.000
MA3	-0.252	-4.696	0.000
SMA1	-0.800	-27.061	0.000

Fuente: Elaboración propia.

Tabla 4.5. Coeficientes y significancia de estos del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$

Coeficiente	Valor	Estadístico t	p-value
AR1	0.940	51.277	0.000
MA1	-0.174	-3.869	0.000
MA2	-0.164	-3.804	0.000
SMA1	-0.817	-27.703	0.000

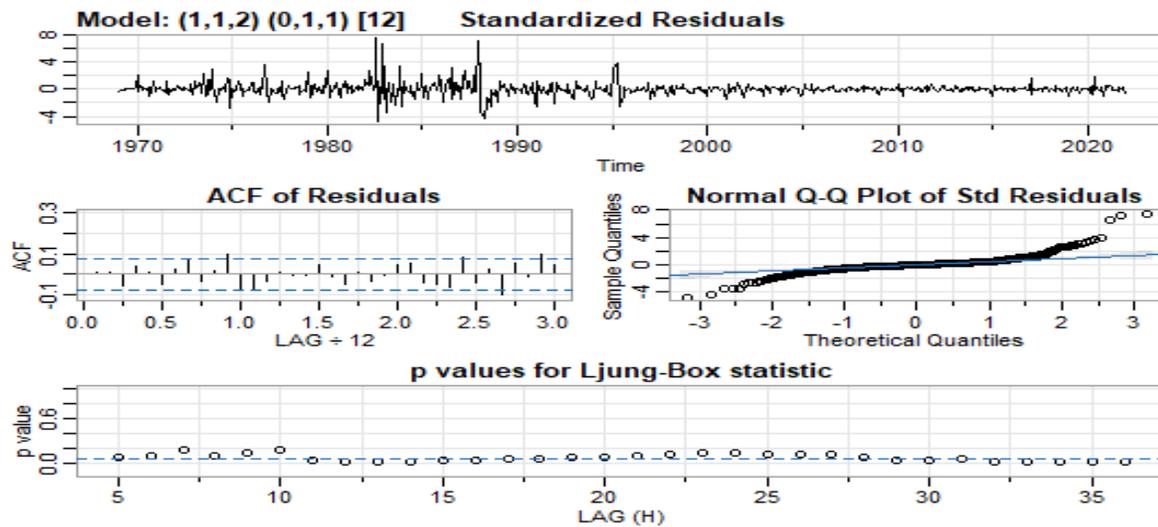
Fuente: Elaboración propia.

Una vez conociendo a fondo los resultados es importante decidir el modelo con el que se llevaran a cabo los pronósticos para el año 2022 y pese a que todos los modelos resultaron ser apropiados, para fines de esta tesis el modelo es aquel con las métricas de error mínimas. El modelo resultante es el modelo $SARIMA(1,1,2)(0,1,1)_{12}$, en adición, vemos que este modelo es capaz de explicar a nuestro conjunto de datos con un menor número de parámetros a diferencia de las demás ecuaciones candidatas con mayor número de estos. Teniendo entonces en cuenta a la ecuación candidata, el paso siguiente es realizar el pronóstico para todo 2022 y después revertir el proceso de la transformación Box-Cox con $\lambda = 0$ para poder conocer los valores en su escala real y poder realizar una comparativa con sus valores observados o al menos poder comparar con aquellos registros disponibles para este 2022.

Para llevar a cabo los pronósticos, se hace uso de todo el conjunto de datos disponible, es decir se emplearon los 636 registros ajustando la ecuación candidata, el comportamiento de los residuos no parece mostrar alertas, la prueba Ljung-Box arroja un valor p de 0.1302 así como también observamos un comportamiento que sigue una distribución Normal de estos. El Gráfico 4.21 muestra el comportamiento de los residuales para esta ecuación tomando en cuenta todo el conjunto de datos disponible.

Una vez hecho el ajuste, los pronósticos para el año 2022 se resumen en la Tabla 4.6 y en la Tabla 4.7 sus valores de pronóstico en la escala real.

Gráfico 4.21. Residuales del modelo $SARIMA(1,1,2)(0,1,1)_{12}$ tomando todo el conjunto de datos disponible.



Fuente: Elaboración propia.

Tabla 4.6. Pronósticos para el año 2022 en escala logarítmica.

Mes	Pronóstico
Enero	4.770
Febrero	4.774
Marzo	4.778
Abril	4.776
Mayo	4.775
Junio	4.778
Julio	4.783
Agosto	4.786
Septiembre	4.790
Octubre	4.796
Noviembre	4.804
Diciembre	4.809

Fuente: Elaboración propia.

Tabla 4.7. Pronósticos para el año 2022 en su escala normal.

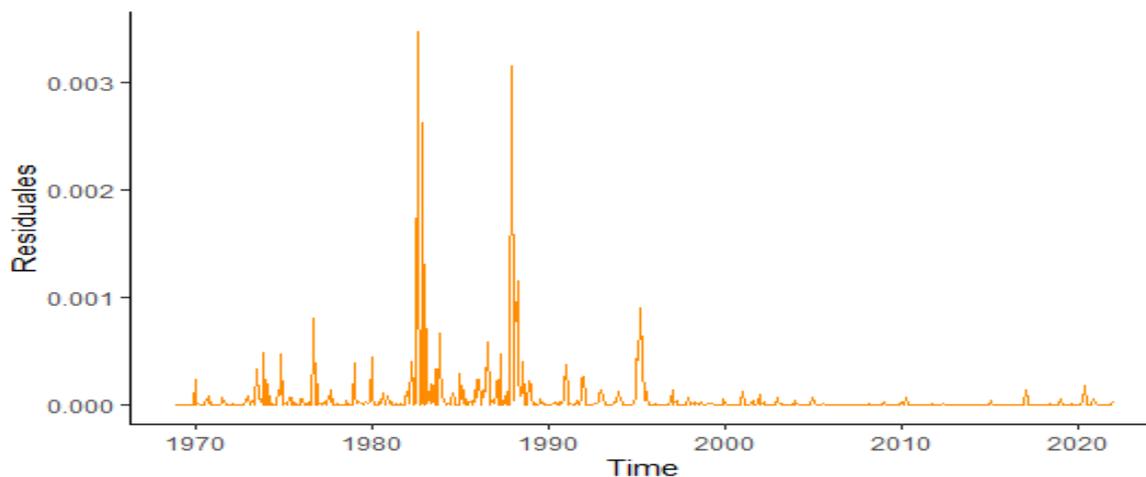
Mes	Pronóstico
Enero	117.916
Febrero	118.367
Marzo	118.849
Abril	118.684
Mayo	118.559
Junio	118.914
Julio	119.427
Agosto	119.785
Septiembre	120.299
Octubre	121.029
Noviembre	121.958
Diciembre	122.567

Fuente: Elaboración propia.

Recuperando el tema de la volatilidad implícita en las series de tiempo económicas, así como también el comportamiento de la economía en pleno 2022 en donde se observaron incrementos bastante agresivos respecto a la inflación debido a los rezagos de la pandemia ocasionada por el virus del COVID 19, así como también por los conflictos bélicos a nivel internacional entre Rusia y Ucrania en donde se vio claramente una afectación directa a la cadena de suministros, se decidió ajustar un modelo GARCH al conjunto de los residuales de la ecuación $SARIMA(1,1,2)(0,1,1)_{12}$. Lo anterior, con la finalidad de construir intervalos de confianza para los pronósticos previamente mostrados, así como también entender el comportamiento de esta serie para aquellos choques de volatilidad que lleguen a impactar al conjunto de datos y tener una perspectiva de pronósticos bajo el escenario más volátil (límite superior del intervalo de confianza) y el menos volátil (límite inferior del intervalo de confianza) en donde como posible interpretación la economía no estaría siendo afectada por los choques de volatilidad.

Para llevar a cabo la construcción de los intervalos de confianza se extrajeron los residuales del modelo y a estos se realizó un tratamiento similar al conjunto de datos original. Esto con la finalidad de determinar el orden de la ecuación GARCH, para ello en primera estancia se elevó al cuadrado dichos residuales para poder ubicar aquellos periodos con mayor volatilidad²⁰ (Gráfico 4.22) dentro de la serie de datos. Posteriormente, se creó el gráfico de la función de autocorrelación para identificar aquellos rezagos significativos y encontrar un orden a la ecuación (Gráfico 4.23).

Gráfico 4.22. Residuales al cuadrado del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$ tomando todo el conjunto de datos disponible.



Fuente: Elaboración propia.

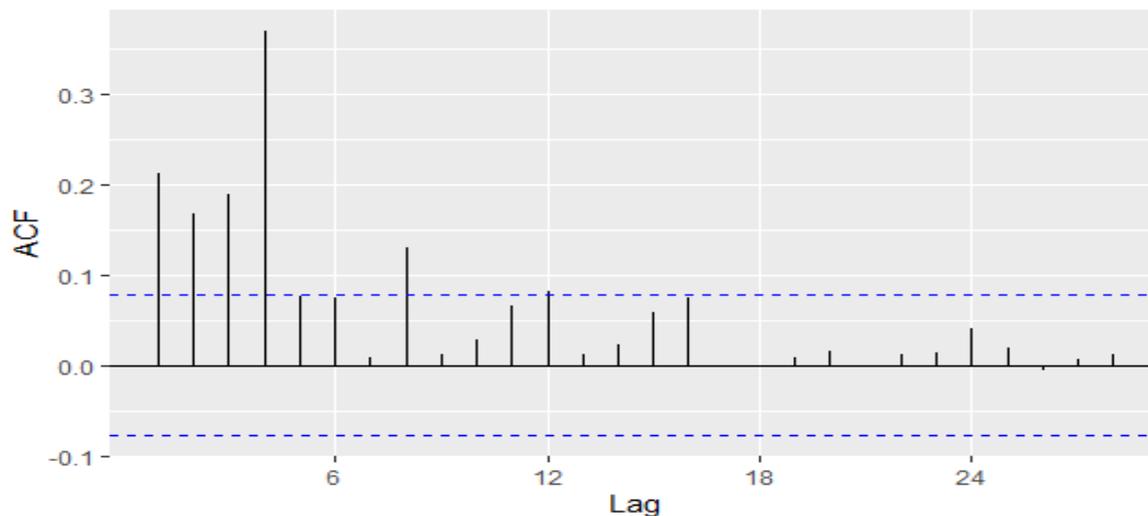
Como puede apreciarse existen periodos marcados por la volatilidad alrededor de 1980 y poco antes del 2000. Estos periodos tienen un respaldo económico cuya explicación se encuentra en el capítulo 3 de esta tesis. Para tener un respaldo estadístico más fuerte respecto a la presencia de volatilidad implícita en la serie se utilizará la prueba ARCH cuyo contraste de hipótesis se explica como sigue:

$$H_0 = \text{La serie de tiempo no presenta efectos ARCH}$$

²⁰ Véase el capítulo 2 en donde se explican a detalle los sucesos económicos que han impactado al desarrollo de la economía mexicana a lo largo de su historia.

Para no rechazar la hipótesis nula el valor p deberá ser mayor a un nivel de significancia, para fines de esta tesis el valor de significancia es $\alpha = 0.05$. Por otro lado, para rechazar la hipótesis nula el valor p deberá ser mucho menor a ese nivel de significancia. Los resultados de esta prueba aplicada al conjunto de residuales de este modelo se resumen en la tabla A.3 en el apartado de Anexos en donde puede apreciarse que el valor p de la prueba resulta ser muy pequeño ($p < 0.05$) en comparación al nivel de significancia, así como también alcanza un valor del estadístico Chi-Cuadrado de 121.370 con 12 grados de libertad. Con lo anterior se tiene evidencia de que la serie presenta efectos ARCH. Una vez se han confirmado los efectos ARCH, el paso siguiente es identificar el orden de la ecuación, esto por medio del gráfico de la función de autocorrelación de los residuales al cuadrado el cual luce como sigue (gráfico 4.23).

Gráfico 4.23. ACF de los Residuales al cuadrado del modelo $SARIMA(1, 1, 2)(0, 1, 1)_{12}$ tomando todo el conjunto de datos disponible.

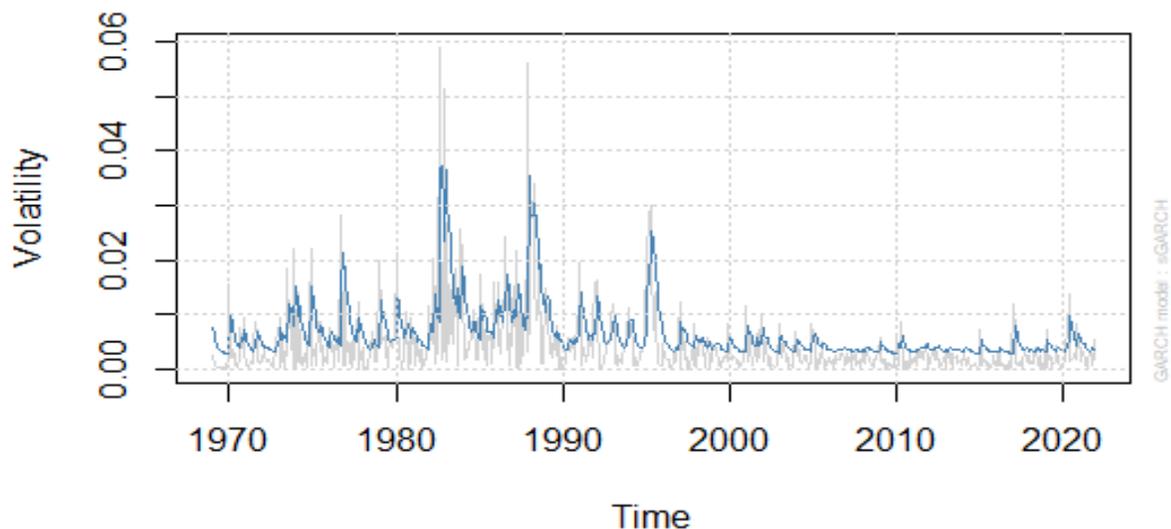


Fuente: Elaboración propia.

Como es posible apreciarse en el gráfico, existe un total de 5 rezagos significativos por lo que bien se puede ajustar una ecuación ARCH (5). Sin embargo, para mantener un menor número de parámetros se optó por ajustar una ecuación GARCH (1,1) dejando que la naturaleza de este modelo explique a través de su varianza pasada al conjunto de los residuales del modelo seleccionado.

Al ajustar la ecuación para el modelo GARCH (1,1) los parámetros de ésta son significativos, así mismo se observa que los resultados se apegan a las propiedades deseadas para ecuaciones de este tipo²¹. En la Tabla A.3 en el apartado de anexos se muestra un resumen de los parámetros del modelo GARCH (1,1), en donde es posible apreciar que cada coeficiente que acompaña a la ecuación resulta ser significativo por los valores obtenidos para el estadístico t (3.563, 6.871, 31.986) y con valores p menores a 0.05, así mismo en el Gráfico 4.24 se muestra el ajuste del modelo respecto al valor real que toma cada residual.

Gráfico 4.24. Ajuste de la ecuación GARCH (1,1) a los residuales de la ecuación candidata.



Fuente: Elaboración propia.

Como puede apreciarse el ajuste es bueno, pues el modelo replica en buena manera aquellos comportamientos en donde existe mayor volatilidad, así mismo la variación aprendida se ajusta muy bien para aquellos periodos en donde existe mayor estabilidad dentro de la serie. El paso siguiente consistió en pronosticar el año 2022 la volatilidad de la serie para posteriormente ajustar intervalos de confianza a las predicciones. El principal motivo de la construcción de intervalos de confianza radica en poder tener una expectativa

²¹ Véase el Capítulo 3, en donde en la metodología de cada modelo se explican a detalle las propiedades que deben de cumplirse al momento de usar modelos ARCH/GARCH para la volatilidad.

en base a probabilidades sobre el valor que pueda obtener del conjunto de datos en un futuro (Hyndman & Athanasopoulos, 2022). Para generar dichos intervalos de confianza se utiliza la fórmula:

$$\hat{y}_{T+h|T} \pm c\hat{\sigma}_h \quad (1)$$

En donde: $\hat{y}_{T+h|T}$: Es la predicción, en este caso la predicción a 1 año utilizando el modelo $SARIMA(1,1,2)(0,1,1)_{12}$; $\hat{\sigma}_h$: Refiere a la volatilidad pronosticada utilizando el modelo GARCH (1,1). C: Es el multiplicador que depende al grado de probabilidad a utilizar, estos valores en relación con el porcentaje pueden ir desde 50 hasta 99%. El valor que puede tomar c se resume en la Tabla 4.8.

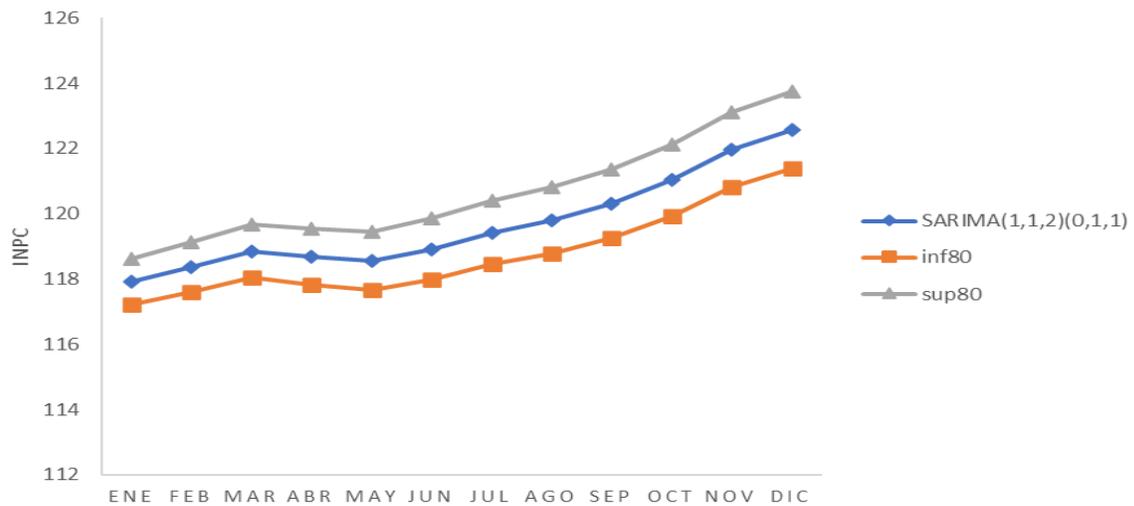
Tabla 4.8. Posibles valores de c según el grado de probabilidad

Porcentaje	Multiplicador
50	0.67
55	0.76
60	0.84
65	0.93
70	1.04
75	1.15
80	1.28
85	1.44
90	1.64
95	1.96
96	2.05
97	2.17
98	2.33
99	2.58

Fuente: (Hyndman & Athanasopoulos, Forecasting: Principles and Practice, 2022)

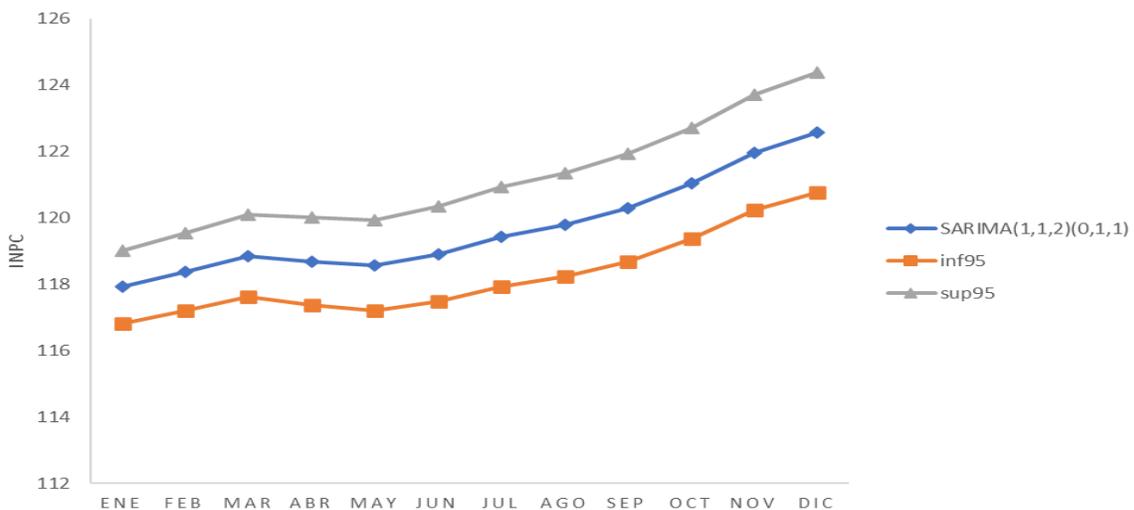
Es común utilizar intervalos de confianza al 80 y 95%, por ello se han de aplicar dichos intervalos a los pronósticos. Los Gráficos 4.25 y 4.26 muestran los resultados de llevar a cabo la formula (1) teniendo en cuenta que los datos han de revertirse dejándose solamente en la escala original de los mismos, así mismo la Tabla 4.9 muestra los resultados numéricos de incluir intervalos de confianza al 80 y 95%.

Gráfico 4.25. Intervalos de confianza al 80% para el pronóstico del INPC en México.



Fuente: Elaboración propia.

Gráfico 4.26. Intervalos de confianza al 95% para el pronóstico del INPC en México.



Fuente: Elaboración propia.

Tabla 4.9. Pronósticos del INPC en México, así como los intervalos de confianza contruidos con (1) al 80 y 95%

Mes	SARIMA (1,1,2) (0,1,1)	inf ₈₀	sup ₈₀	inf ₉₅	sup ₉₅
Enero	117.916	117.203	118.633	116.827	119.015
Febrero	118.368	117.605	119.135	117.201	119.545
Marzo	118.849	118.038	119.665	117.609	120.101
Abril	118.684	117.832	119.543	117.382	120.002
Mayo	118.559	117.667	119.458	117.196	119.938
Junio	118.914	117.980	119.854	117.488	120.357
Julio	119.427	118.452	120.410	117.938	120.935
Agosto	119.785	118.771	120.807	118.236	121.353
Septiembre	120.299	119.247	121.361	118.691	121.929
Octubre	121.029	119.936	122.132	119.360	122.722
Noviembre	121.958	120.824	123.103	120.226	123.716
Diciembre	122.567	121.394	123.750	120.776	124.384

Fuente: Elaboración propia.

Como es posible apreciar, los intervalos al 80% resultan ser un poco más conservadores respecto al pronóstico otorgado por la ecuación $SARIMA(1,1,2)(0,1,1)_{12}$ a diferencia de los intervalos al 95% en donde se observa mayor posibilidad de valores para el INPC en México. Al implementar estos intervalos se logra una amplia expectativa sobre los diferentes escenarios que pudiesen ocurrir dentro de la economía mexicana, con lo cual se logra un escenario en donde el ritmo del INPC crece lentamente. Es decir, el aumento de la canasta básica sería moderado apegándose a la tendencia que estos valores tienen en un inicio, lo anterior haciendo referencia al límite inferior del intervalo, así mismo es posible tener un escenario completamente distinto en el cual el ritmo de crecimiento de la canasta básica es más agresivo y por ende se esperaría que el INPC crezca a un ritmo mayor, lo que bien puede traducirse como periodos en donde existen efectos de volatilidad dentro de la economía mexicana y con esto se hace referencia al límite superior del intervalo de confianza.

Por otra parte, existe el pronóstico que por sí solo otorga el modelo candidato, con este el ritmo de crecimiento se encuentra justo por la mitad de los intervalos de confianza previamente mencionados. El implementar un modelo que entienda la volatilidad ayuda mucho sobre todo cuando el objetivo de pronóstico es un instrumento importante dentro de la economía ya que con esto se considera en buena medida el comportamiento de aquellos sucesos que puedan impactar en el ciclo de vida económico, como lo es el caso de la actual guerra entre Rusia y Ucrania, así como también la pandemia del COVID-19 que fueron sucesos completamente aleatorios que tuvieron a bien impactar negativamente dentro de la economía mexicana haciendo que los niveles inflacionarios crezcan a un ritmo que jamás en la historia mexicana se ha visto.

Para medir el comportamiento de la inflación con base en los pronósticos se recupera la fórmula de razón de cambio que utiliza el INEGI para saber el incremento de la inflación de un periodo a otro. Al comparar la variación entre los índices de precios proyectados puede tenerse un acercamiento real al momento de tratar de medir la inflación en México. Para cumplir con el objetivo anteriormente mencionado se aplicó la siguiente fórmula:

$$\left(\frac{INPC_t}{INPC_{t-1}} - 1 \right) * 100 \quad (2)$$

En donde: $INPC_t$: Representa el valor del INPC en el tiempo t, dicho tiempo puede estar expresado de manera trimestral, mensual o anual. $INPC_{t-1}$: Representa el valor del INPC en un periodo anterior al tiempo t, de igual modo este tiempo puede estar expresado de manera trimestral, mensual o anual.

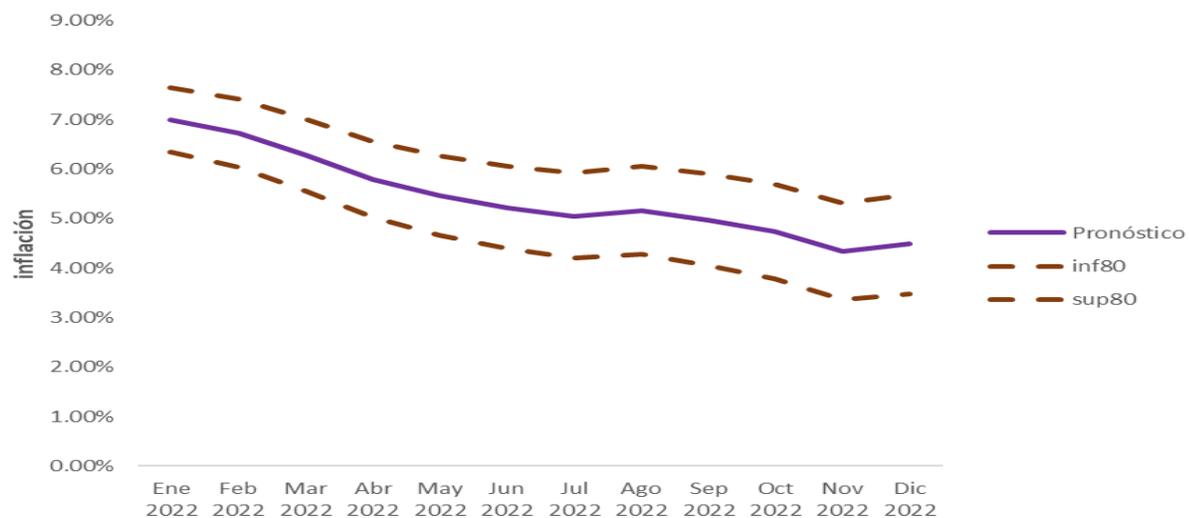
Para poder representar el cómo se ha comportado la inflación en México de manera anual, se recuperan los pronósticos de todo 2022 y con la ayuda de (2) se compara la variación (incremento o decremento) respecto al nivel del INPC registrado en 2021 por el INEGI. La Tabla 4.10 muestra la inflación mensual anualizada con sus respectivos intervalos de confianza al 80 y 95%, así como los Gráficos 4.27 y 4.28 muestran los niveles de inflación proyectada con intervalos de confianza del 80 y 95% respectivamente.

Tabla 4.10. Pronósticos de la inflación mensual anualizada usando (2) en México, así como los intervalos de confianza contruidos con (1) al 80 y 95%

Fecha	Pronóstico	inf ₈₀	sup ₈₀	inf ₉₅	sup ₉₅
Enero 2022	6.99%	6.35%	7.64%	6.00%	7.99%
Febrero 2022	6.73%	6.04%	7.42%	5.68%	7.79%
Marzo 2022	6.28%	5.56%	7.01%	5.17%	7.40%
Abril 2022	5.79%	5.03%	6.55%	4.63%	6.96%
Mayo 2022	5.46%	4.67%	6.26%	4.25%	6.69%
Junio 2022	5.22%	4.39%	6.05%	3.95%	6.49%
Julio 2022	5.05%	4.20%	5.92%	3.74%	6.38%
Agosto 2022	5.17%	4.28%	6.06%	3.81%	6.54%
Septiembre 2022	4.97%	4.05%	5.90%	3.57%	6.39%
Octubre 2022	4.73%	3.79%	5.69%	3.29%	6.20%
Noviembre 2022	4.34%	3.37%	5.32%	2.86%	5.84%
Diciembre 2022	4.48%	3.48%	5.49%	2.96%	6.03%

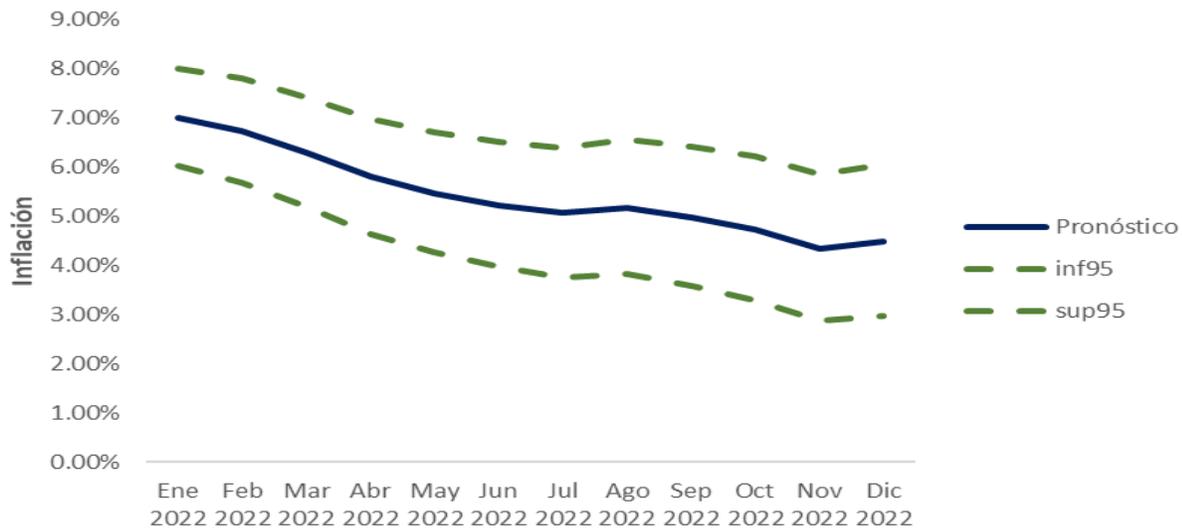
Fuente: Elaboración propia.

Gráfico 4.27. Intervalos de confianza al 80% para el pronóstico de la inflación en México usando (2) y (1).



Fuente: Elaboración propia.

Gráfico 4.28. Intervalos de confianza al 95% para el pronóstico de la inflación en México usando (2) y (1).



Fuente: Elaboración propia.

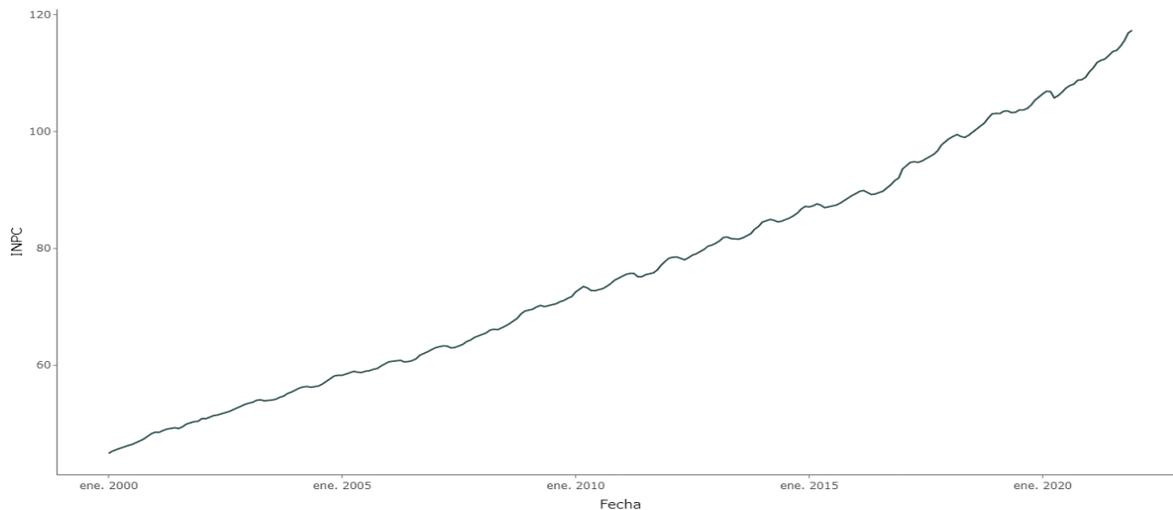
Con la implementación de lo anterior se observa que los pronósticos sobrepasan el límite objetivo de BANXICO llegando a niveles que oscilan entre el 3.37% como mínimo y un 7.64% para el caso de los intervalos de confianza al 80% y de igual modo niveles comprendidos en el 2.89% a 7.99% para aquellos intervalos construidos en un 95%. Para concluir la implementación de este método de pronóstico se observa que en resumidas cuentas el nivel esperado de la inflación (o al menos lo que pronostica el modelo $SARIMA(1,1,2)(0,1,1)_{12}$ con intervalos de confianza construidos por un modelo $GARCH(1,1)$) es relativamente alto o al menos para los niveles de control impuestos por BANXICO. De lo anterior se desprende que al menos para fines de esta tesis el modelo logra adaptarse a los pronósticos pese a las duras condiciones económicas que se logran ver para este 2022.

4.2 Modelo SARIMAX

Para el último modelo se modifica el periodo de tiempo debido a la disponibilidad de los datos de las variables exógenas contempladas, la temporalidad inicia en enero del año 2000 a diciembre de 2021 tomando en cuenta un total de 264 datos.

Como se había detallado antes, el comportamiento del INPC a partir de los noventa es linealmente creciente con algunas variaciones, este modelo no tomará en cuenta la volatilidad de la serie.

Gráfico 4.29. Índice Nacional de Precios al Consumidor: enero de 2000 a diciembre de 2021.

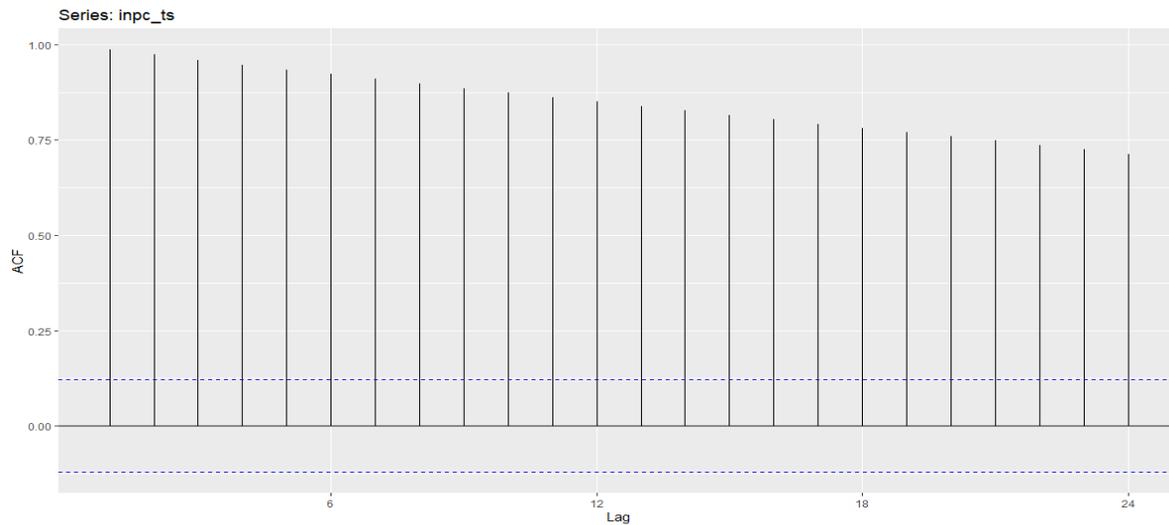


Fuente: Elaboración propia.

Como indica la metodología, se realizan las pruebas de bondad de ajuste al proceso para verificar si cumple con el supuesto de la estacionariedad en media. Los resultados Kwiatkowski-Phillips-Schmidt-Shin ($q = 4.50$, $p < 0.05$) y Dickey Fuller ($q = 2.284$, $p < 0.1$), según las hipótesis manejadas de cada prueba definidas anteriormente, el valor p obtenido es indicativo de que se rechaza que el proceso sea estacionario en media, el resultado de la prueba se puede corroborar en la tabla número A.4 del apartado de anexos.

Este resultado se respalda con el autocorrelograma de la serie (Gráfico 4.30) donde se ve a todos los rezagos fuera de la banda de confianza y a la mayoría cerca de uno, además de que el decaimiento de la serie es lento, pero un poco menos que para el proceso del modelo 1 donde se contempla una temporalidad más larga.

Gráfico 4.30. Función de Autocorrelación del Índice Nacional de Precios al Consumidor: 2000-2021.

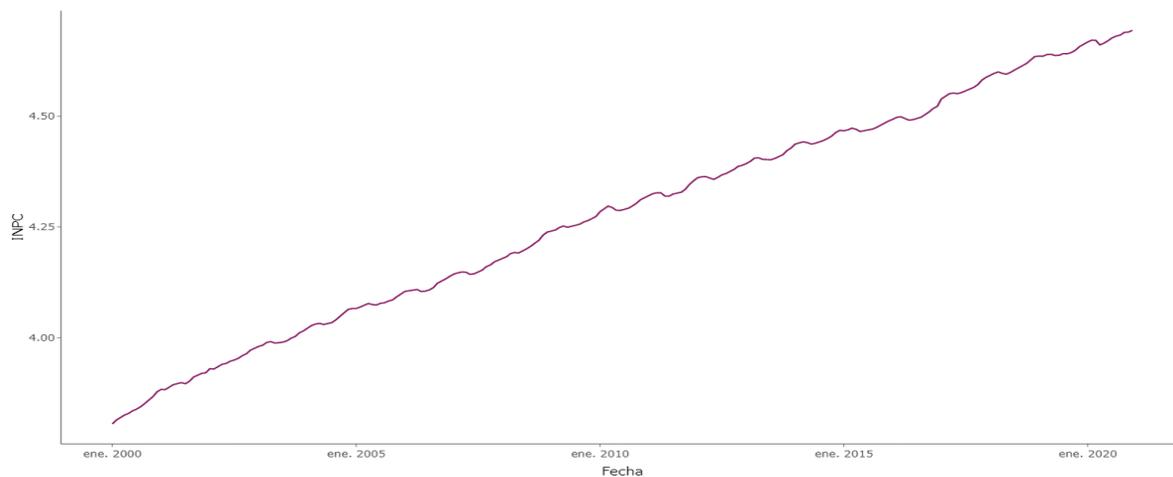


Fuente: Elaboración propia.

De la misma forma que para el modelo 1, para comenzar con la construcción del modelo ARIMAX los datos del INPC se dividen en un conjunto de entrenamiento que se usará para ajustar los parámetros de la ecuación y otro de prueba para medir la precisión de los modelos, la división de cada conjunto va de acuerdo con el horizonte temporal que se desea pronosticar, para este caso, el conjunto de entrenamiento irá de enero del año 2000 a diciembre de 2020 y el conjunto de prueba de enero a diciembre de 2021.

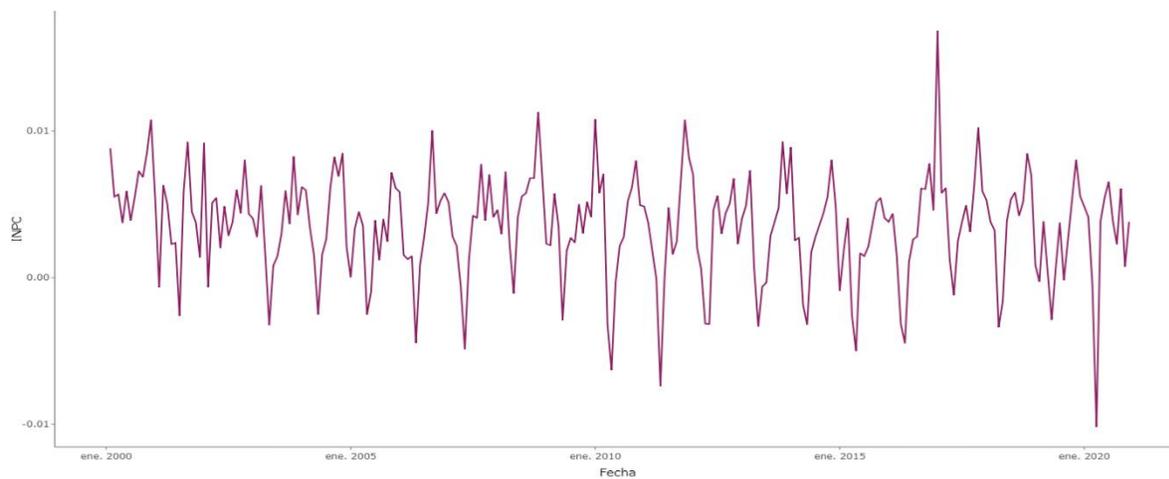
El conjunto de entrenamiento tiene el mismo problema de tendencia debido a que el INPC ha ido en aumento con el paso del tiempo. Entonces, y para corregir el problema de estacionariedad se usa una transformación logarítmica y se realizan primeras diferencias.

Gráfico 4.31. INPC conjunto de entrenamiento con transformación logarítmica.



Fuente: Elaboración propia.

Gráfico 4.32. INPC conjunto de entrenamiento transformado con primeras diferencias.



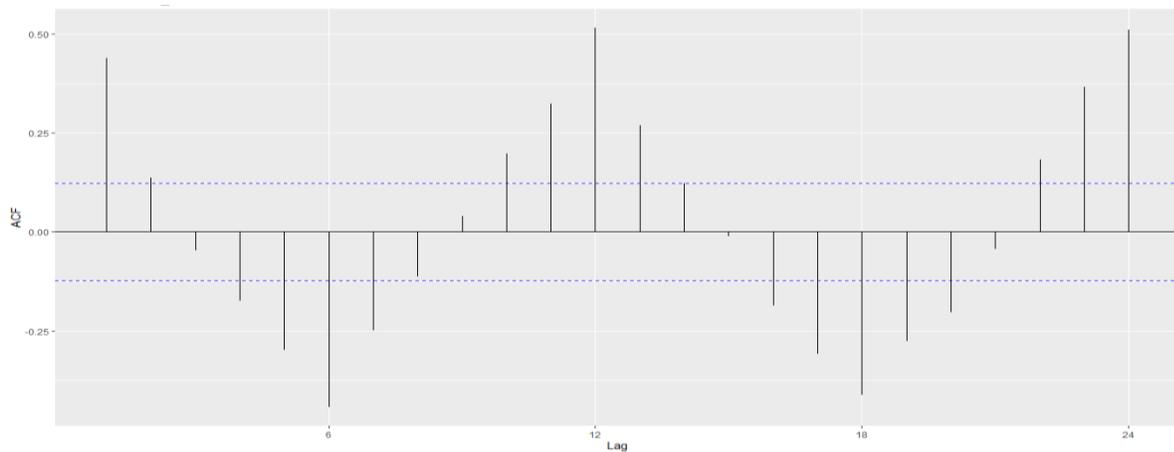
Fuente: Elaboración propia.

Con el tratamiento de los datos visualmente se aprecia que se ha corregido el problema de reversión a la media (Gráfico 4.32), ya que con la reducción de la temporalidad se omitieron los movimientos más bruscos que trajeron las crisis económicas en México.

Para confirmar estadísticamente el resultado, se realizan las pruebas de raíz unitaria Kwiatkowski-Phillips-Schmidt-Shin ($q = 0.122$, $p > 0.05$) y Dickey Fuller ($q = -8.459$, $p > 0.1$), el conjunto de entrenamiento transformado y con primeras diferencias ya es estacionario en media (tabla A.5 del apartado de anexos).

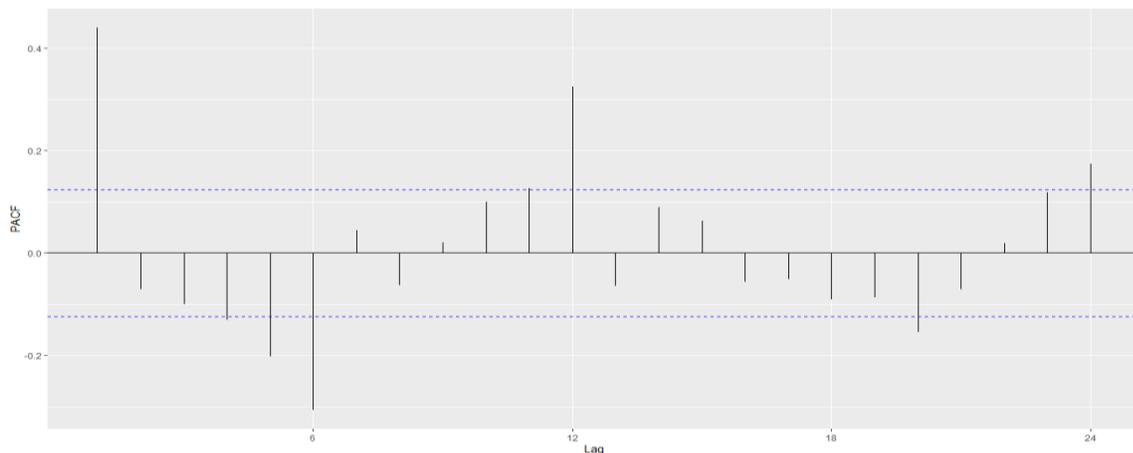
La serie acortada debería presentar problemas de estacionalidad de igual forma que sucedió con los datos completos, para confirmarlo se analizan el autocorrelograma y el autocorrelograma parcial (gráfico 4.33).

Gráfico 4.33. ACF del conjunto de entrenamiento transformado y con primeras diferencias.



Fuente: Elaboración propia.

Gráfico 4.34. PACF del conjunto de entrenamiento transformado y con primeras diferencias.

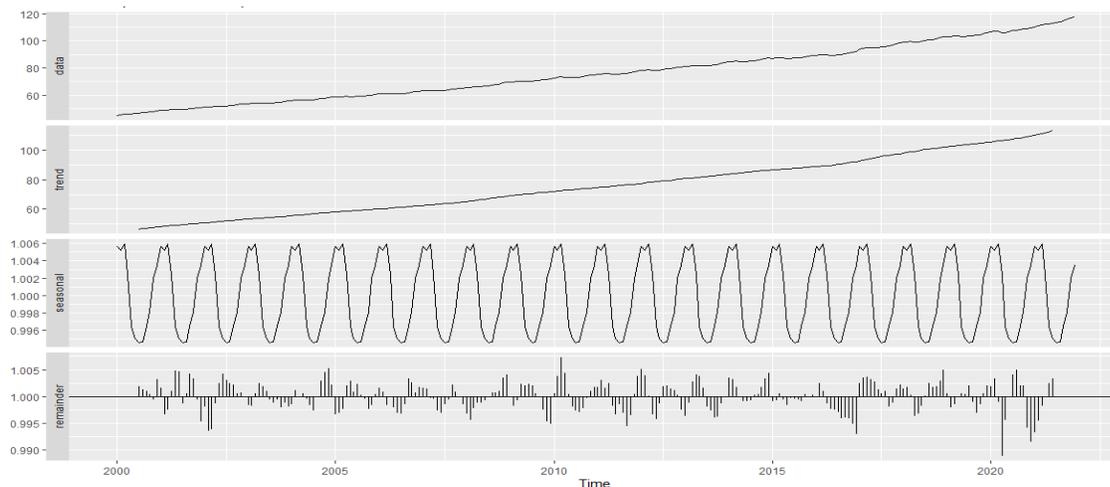


Fuente: Elaboración propia.

El comportamiento de la serie mostrado en el ACF y el PACF pese a ser estacionario no es el apropiado para formular una ecuación, ya que es posible identificar un patrón en los rezagos y un retardo significativo en cada múltiplo de 12 por lo que se procede a hacer un análisis bajo el supuesto antes confirmado de la estacionalidad.

Se realiza la descomposición del conjunto de entrenamiento transformado sin diferencias con un enfoque multiplicativo para identificar y verificar una vez más la presencia de la estacionalidad que se contemplará en las futuras ecuaciones (gráfico 4.35).

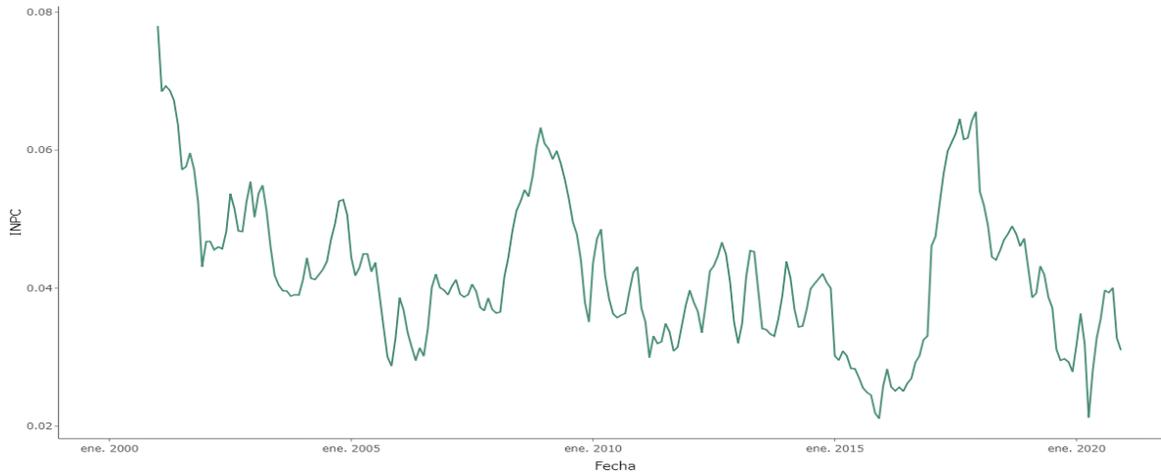
Gráfico 4.35 Descomposición del conjunto de entrenamiento transformado.



Fuente: Elaboración propia.

El gráfico de la descomposición confirma también la tendencia positiva de la serie, y la presencia de la estacionalidad, por lo que se realizan solo diferencias estacionales al proceso para analizar su comportamiento (gráfico 4.36).

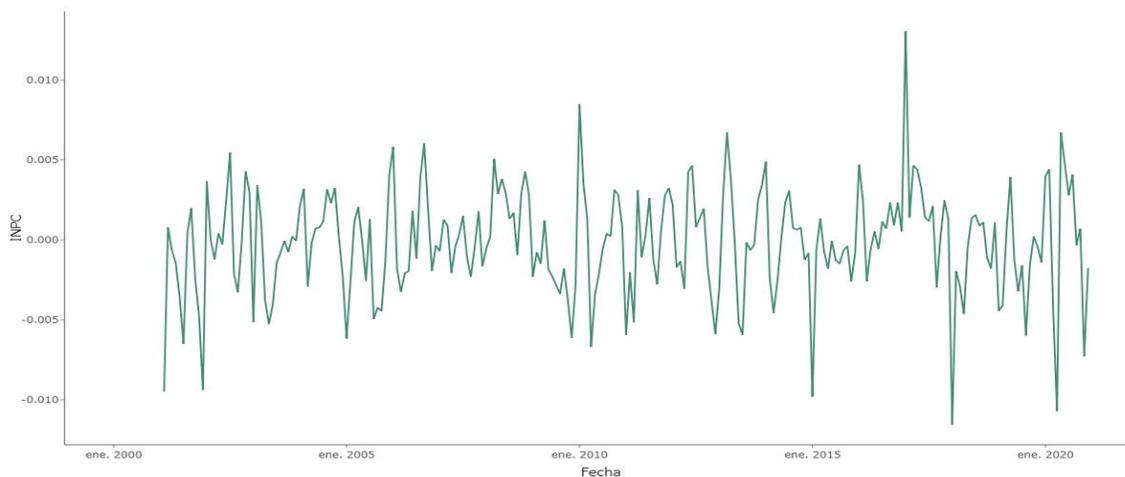
Gráfico 4.36. INPC conjunto de entrenamiento transformado añadiendo diferencias estacionales.



Fuente: Elaboración propia.

El conjunto de entrenamiento transformado y únicamente diferenciado estacionalmente ayuda al proceso con la reversión de la media alrededor del cero, sin embargo, se aprecia que el comportamiento necesita de las primeras diferencias ordinarias para que sea completamente estacionario (gráfico 4.37).

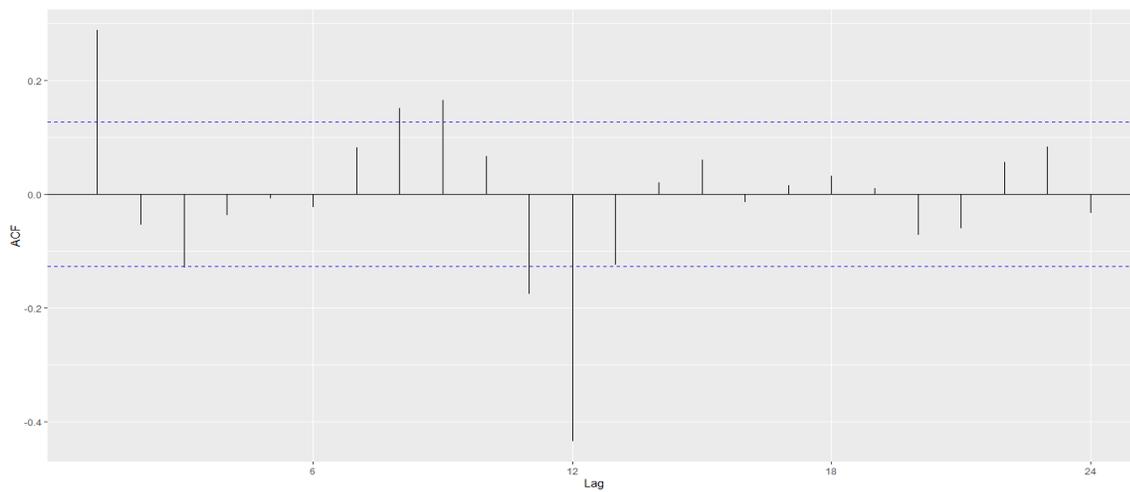
Gráfico 4.37. INPC conjunto de entrenamiento transformado añadiendo diferencias estacionales y primeras diferencias.



Fuente: Elaboración propia.

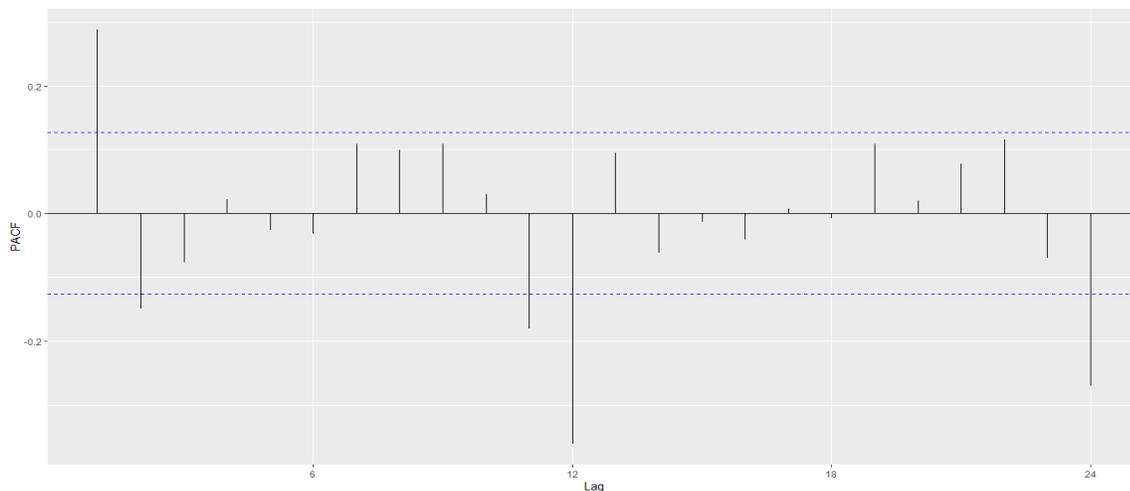
El conjunto de prueba transformado con primeras diferencias ordinarias y estacionales muestra un comportamiento más estable alrededor de cero es similar al proceso con solo diferencias ordinarias, pero se alcanza a percibir un ajuste de los datos, el proceso diferenciado completamente pasa las pruebas de estacionariedad. Los gráficos 4.38 y 4.39 muestran el autocorrelograma y autocorrelograma parcial de la serie con diferencias completas donde se puede apreciar la disminución del patrón de los rezagos significativos.

Gráfico 4.38. ACF del conjunto de prueba transformado y diferenciado completamente.



Fuente: Elaboración propia.

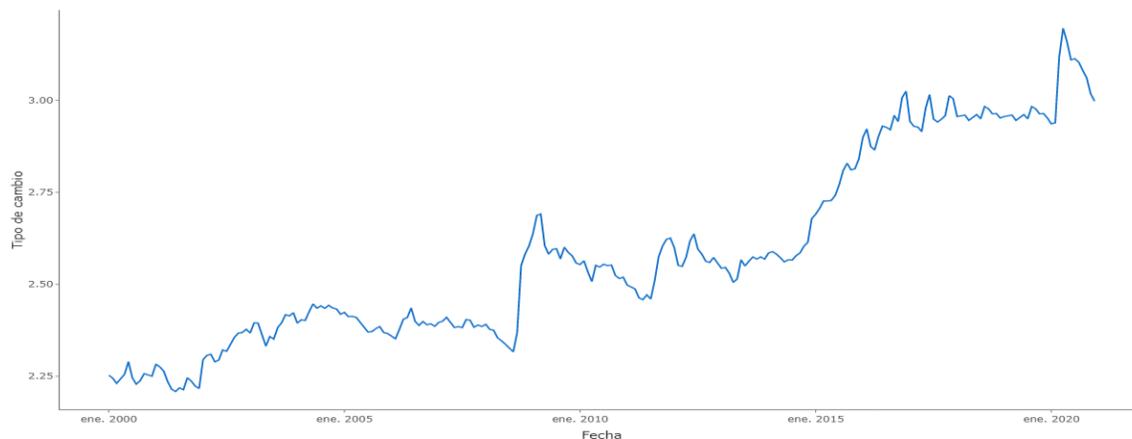
Gráfico 4.39. PACF del conjunto de entrenamiento transformado y diferenciado completamente.



Fuente: Elaboración propia.

Con la serie ahora estacionaria se pueden ajustar las ecuaciones para el modelo basándose en el ACF y PACF, sin embargo, para el enfoque del modelo SARIMAX las variables exógenas que acompañen a la variable independiente, el INPC para este caso, deben ser también series de tiempo estacionarias en media, por lo que a cada variable se le aplicará una transformación logarítmica, primeras diferencias y diferencias estacionales. Para analizar su comportamiento cada variable exógena se divide también en conjunto de prueba y de entrenamiento con la misma temporalidad descrito anteriormente (gráfico 4.40).

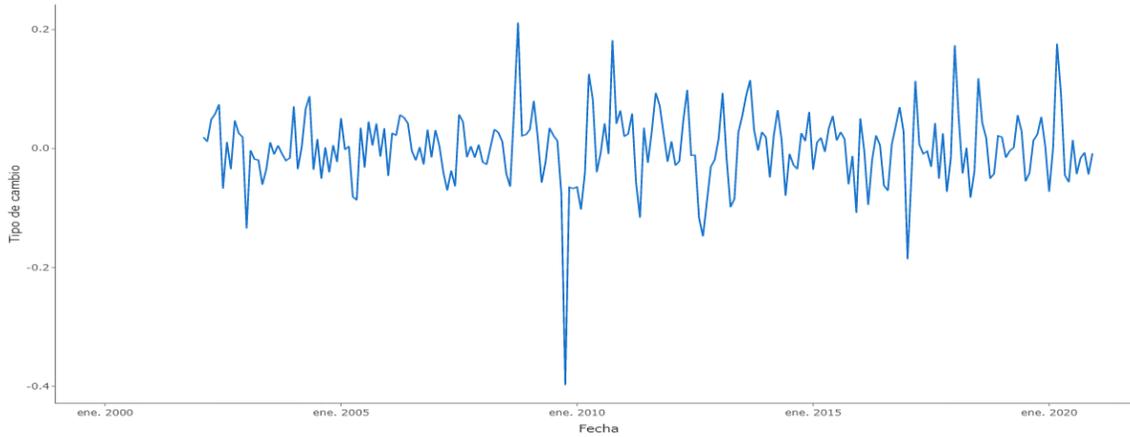
Gráfico 4.40. Comportamiento del conjunto de entrenamiento transformado de la variable Tipo de cambio.



Fuente: Elaboración propia.

El tipo de cambio muestra un comportamiento creciente a pesar de tener variaciones fuertes a lo largo del tiempo, a simple vista se aprecia que no es estacionario en media.

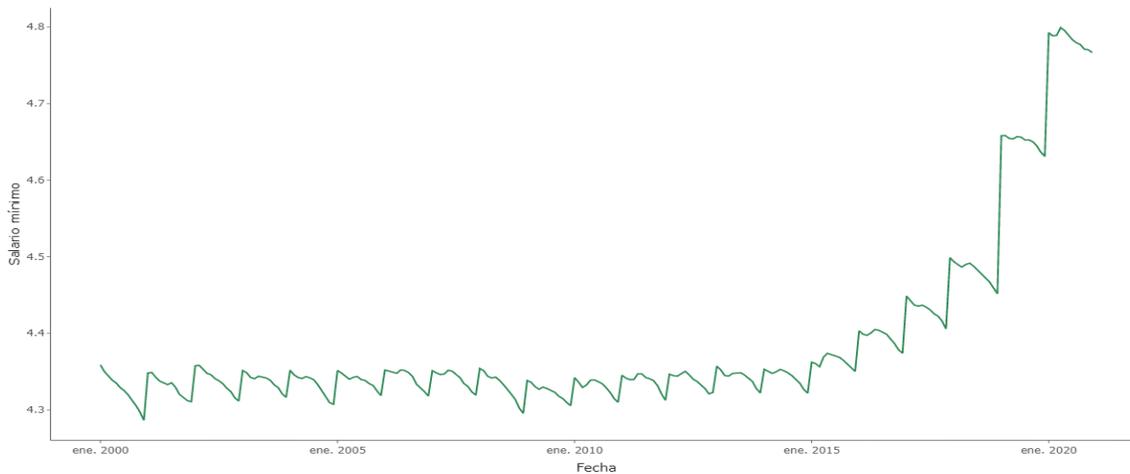
Gráfico 4.41. Primeras diferencias y diferencias estacionales al conjunto transformado de entrenamiento de la variable Tipo de cambio.



Fuente: Elaboración propia.

La serie transformada y con diferencias completas parece visualmente estacionaria.

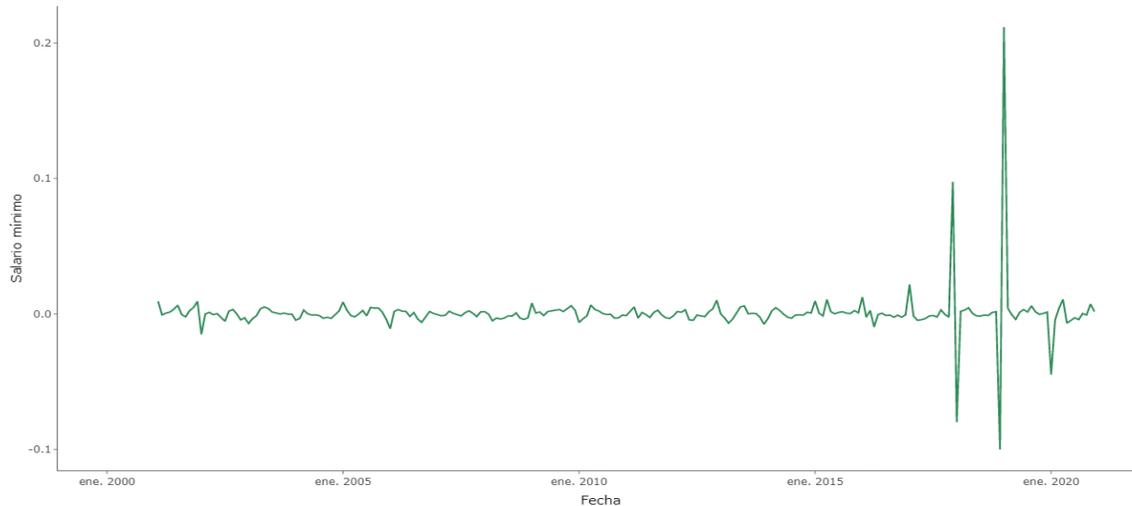
Gráfico 4.42. Comportamiento del conjunto de entrenamiento transformado de la variable Salario mínimo.



Fuente: Elaboración propia.

El Salario mínimo en México se mantuvo casi constante hasta el 2015 donde comienza a incrementarse, fue hasta 2019 y 2020 que se muestra un alza significativa de su valor (gráfico 4.43).

Gráfico 4.43. Primeras diferencias y diferencias estacionales al conjunto de entrenamiento transformado de la variable Salario mínimo.

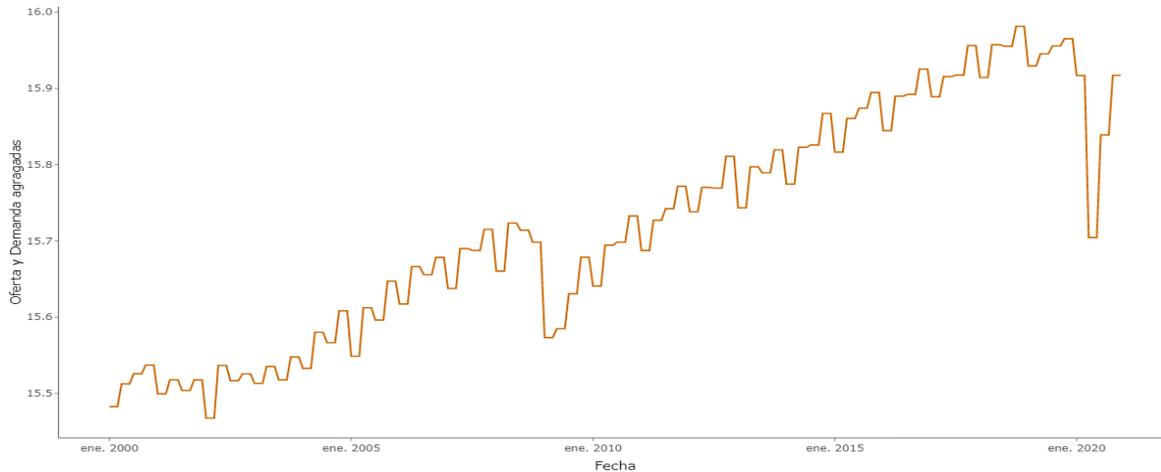


Fuente: Elaboración propia.

La serie de tiempo del salario mínimo mostraba un problema de estacionalidad pues al aplicar únicamente primeras diferencias el proceso no pasaba las pruebas de estacionariedad, agregando las diferencias estacionales se aprecia que se vuelve estacionario en media.

La variable de expectativas inflacionarias se transformó en una variable categórica que toma el valor de 1 si el valor pronosticado por los expertos es mayor que el valor real de la inflación, y 0 en el caso contrario. Al convertirse en una variable categórica la serie de tiempo pasa la prueba estadística de estacionariedad.

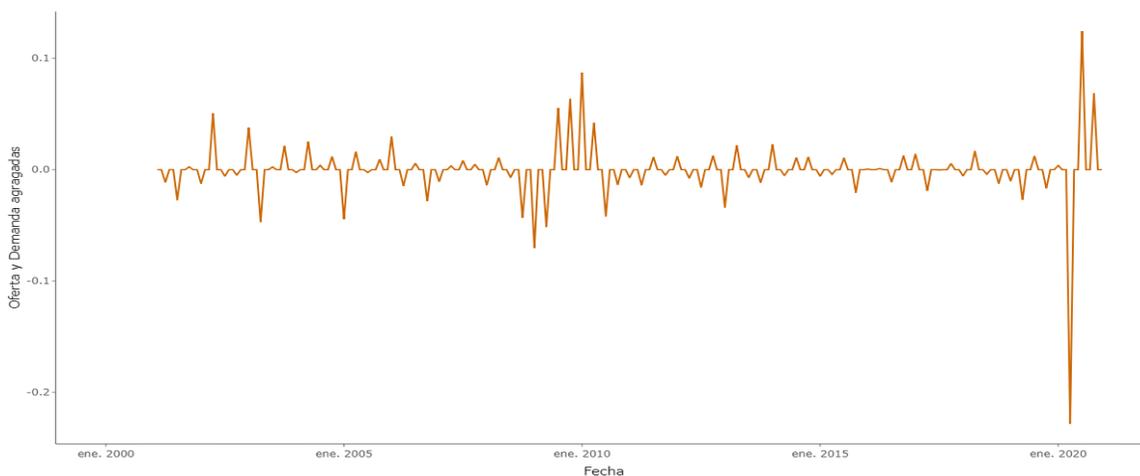
Gráfico 4.44. Comportamiento del conjunto de entrenamiento transformado de la variable Oferta y Demanda agregadas.



Fuente: Elaboración propia.

La variable oferta y demanda agregada en las bases del INEGI se registra de manera trimestral, por tal motivo se dividió el valor reportado entre tres para coincidir con la temporalidad mensual del resto de las variables. Su comportamiento fue también en su mayoría creciente, en 2019 sufrió una fuerte caída.

Gráfico 4.45. Conjunto de prueba transformado de la Variable Oferta y Demanda agregadas con primeras diferencias y diferencias estacionales.



Fuente: Elaboración propia.

Con la transformación de los datos la serie es visiblemente estacionaria. Para confirmar que las variables sean estacionarias en media se realizan las pruebas de raíz unitarias

(tabla A.6 del apartado de anexos). Según los resultados existe evidencia estadística para afirmar que todas las variables exógenas son estacionarias en media.

Los criterios que se tomaron en cuenta en este modelo para clasificar a una ecuación como buena para realizar un pronóstico además de las utilizadas en el modelo convencional SARIMA (coeficientes de la ecuación diferentes de cero, independencia de los residuales, criterio de Akaike y medidas de error pequeñas) serán buscar que el valor p de los coeficientes sea menor a 0.05 y el valor T mayor a 2 en valor absoluto para que cada parámetro y sobre todo el correspondiente a cada variable exógena sea significativa al modelo.

Una vez que las variables exógenas están transformadas en series de tiempo estacionarias pueden agregarse a una matriz y formular la ecuación del modelo SARIMAX. Con base en los Gráficos 4.38 y 4.39 de las Función de Autocorrelación y Funciones de Autocorrelación Parciales del conjunto de entrenamiento transformado y diferenciado completamente se identifica el número de parámetros autorregresivos, de medias móviles, así como los respectivos estacionales: las mejores ecuaciones se muestran a continuación.

En primer lugar, la ecuación $SARIMAX(3,1,1)(1,1,1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias, el desempeño de la ecuación tiene buenos resultados en cuanto a los residuales dado que todos se mantienen alejados de la banda de confianza. La independencia se alcanza rápidamente la prueba Ljung box tiene un valor p de 0.742, las tres variables exógenas tienen muy buena aceptación en el modelo ya que su valor es casi de cero. El coeficiente estacional tiene un valor p de 0.0428 cerca del valor aceptado, sin embargo, su valor t es de -2.0373 lo cual es aceptable para este coeficiente. Los resultados de la ecuación se muestran en el Gráfico 4.46.

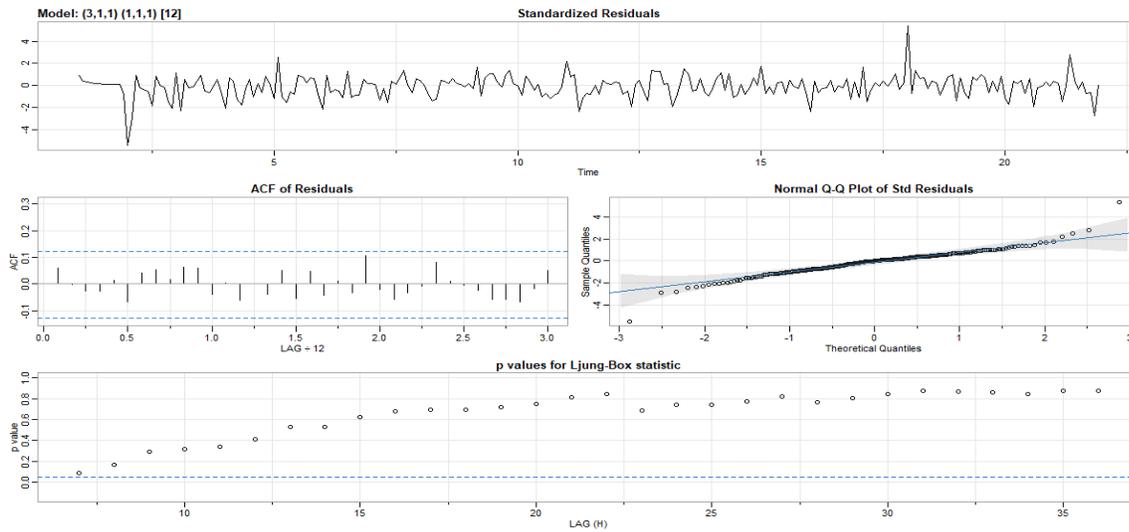
Derivada de la ecuación anterior se decide incrementar un componente autorregresivo, uno de medias móviles y eliminar el autorregresivo correspondiente estacional, resultando

la ecuación $SARIMAX(4,1,3)(0,1,1)_{12}$, con las variables exógenas salario mínimo, oferta y demanda agregada y expectativas inflacionarias, el resultado es similar al obtenido anteriormente para el caso de los residuales que se mantienen todos dentro de la banda de confianza. El valor de la prueba de independencia es igual a 0.5855, todos los coeficientes en esta ecuación tienen coeficientes diferentes de cero y resultan significativos con valores p de prácticamente 0. Los resultados se encuentran en el Gráfico 4.47.

Lo siguiente fue encontrar una ecuación más sencilla por lo que se construyó el modelo $SARIMAX(3,1,1)(0,1,1)_{12}$ regresando a 3 autorregresivos y eliminando el estacional que no era significativo para el modelo. Se usaron las variables salario mínimo, oferta y demanda agregada y las expectativas inflacionarias, el resultado fue que los valores p del parámetro autoregresivo número 3 y de la variable exógena salario mínimo incrementaron, pero se siguen manteniendo por debajo de 0.05, el valor t es mayor a 2 en valor absoluto por lo que serán considerados valores aceptables. El estadístico de independencia Ljung box es igual a 0.6405, en los residuales podrían decirse que hay algunos retardos significativos, sin embargo, están dentro de la banda de confianza, los resultados se muestran en el Gráfico 4.48.

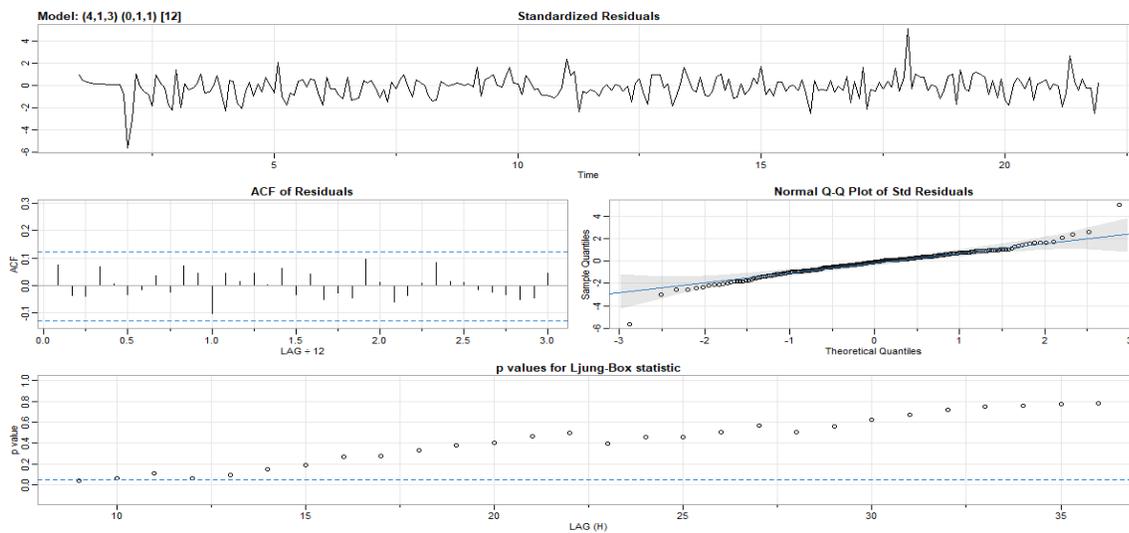
Continuando con la ecuación $SARIMAX(3,1,1)(0,1,1)_{12}$ se probó cambiar las variables exógenas, incluyendo solamente el salario mínimo y la oferta y demanda agregadas, debido a que la variable *expectativas inflacionarias* ha tenido en los resultados anteriores, un coeficiente diferente de cero pero muy cercano a cero. Como resultado se eliminó uno de los rezagos significativos mostrados en el Gráfico 4.48. La independencia se alcanza rápidamente y el resultado de la prueba es igual a 0.7557, y al igual que en el modelo anterior se incrementa el valor p, pero se considera dentro de los parámetros normales. Los resultados de los residuales se muestran en el Gráfico 4.49.

Gráfico 4.46. Residuales del modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



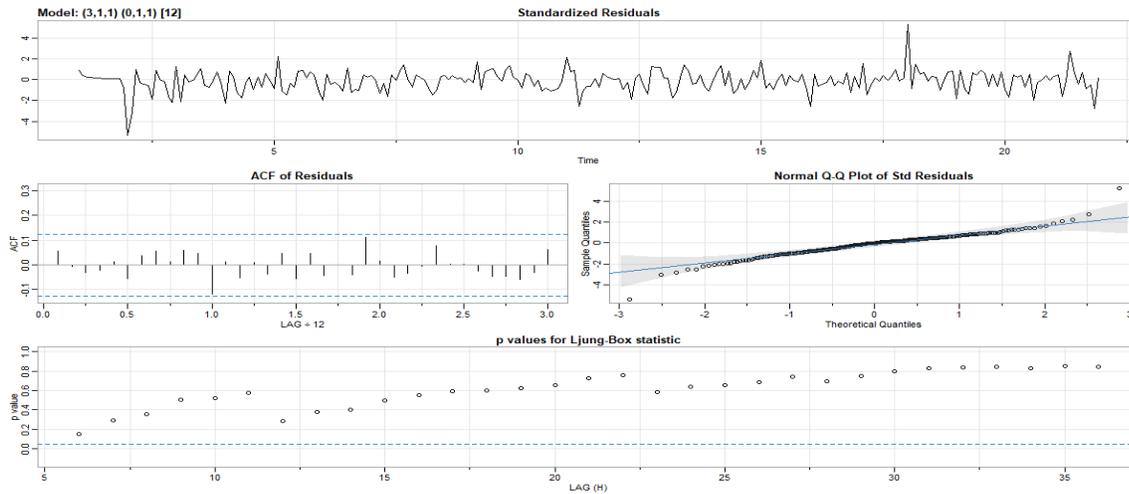
Fuente: Elaboración propia.

Gráfico 4.47. Residuales del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



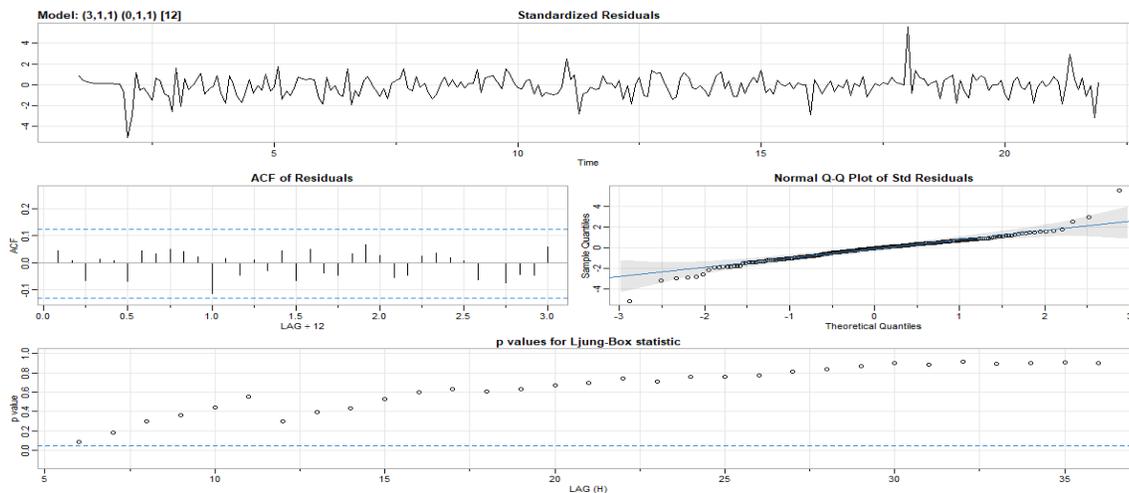
Fuente: Elaboración propia.

Gráfico 4.48. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



Fuente: Elaboración propia.

Gráfico 4.49. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con variables salario mínimo y oferta y demanda agregada.

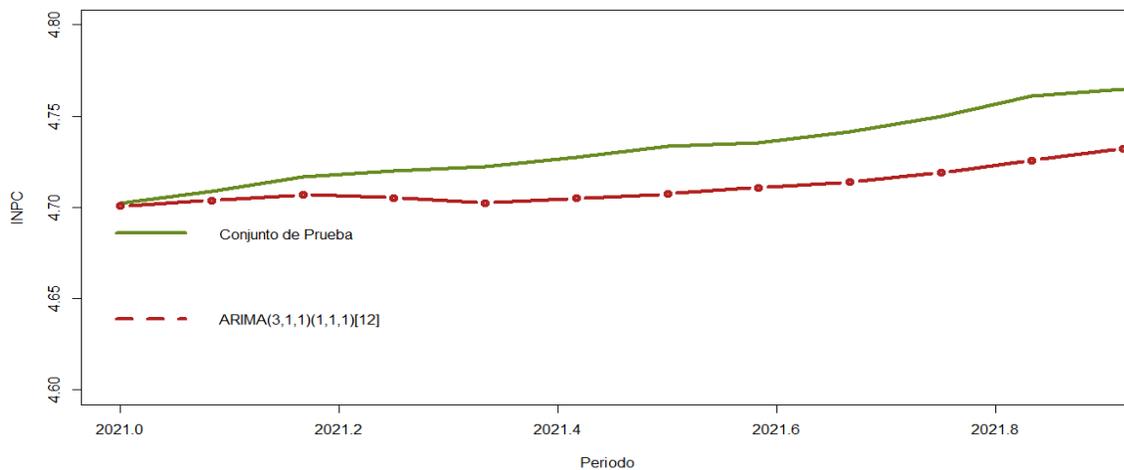


Fuente: Elaboración propia.

A continuación, se mide la precisión de los cuatro modelos utilizando el conjunto de prueba pronosticando el año 2021 y obteniendo las medidas de error RMSE, MAE y MAPE aceptando como bueno al modelo que minimice estos errores. Los resultados de la

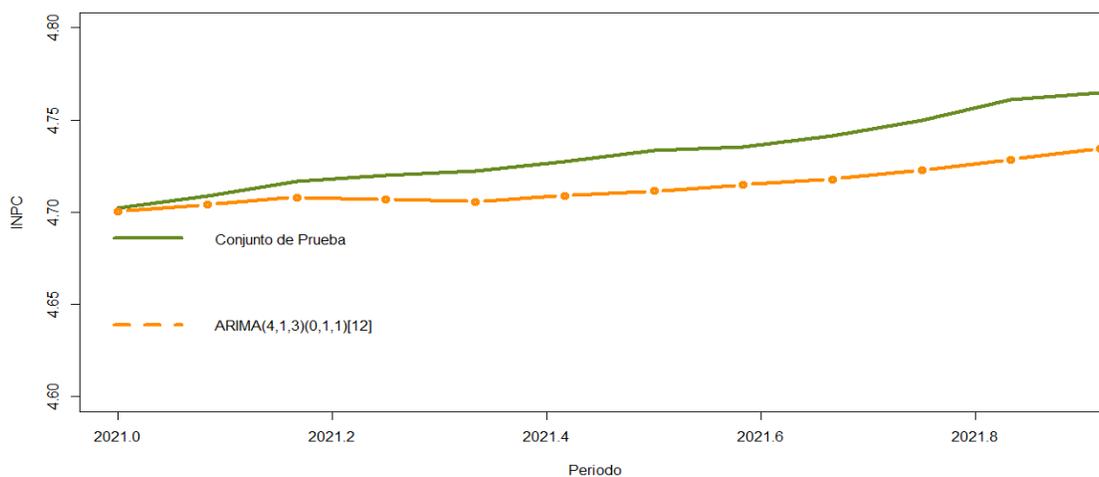
comparación gráfica, así como la de las medidas estadísticas y la descripción de la significancia de los parámetros se encuentran en los Gráficos 4.50, 4.51, 4.52, 4.53 y las Tablas 4.11, 4.12, 4.13, 4.14 respectivamente.

Gráfico 4.50. Comparación del conjunto de prueba vs modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



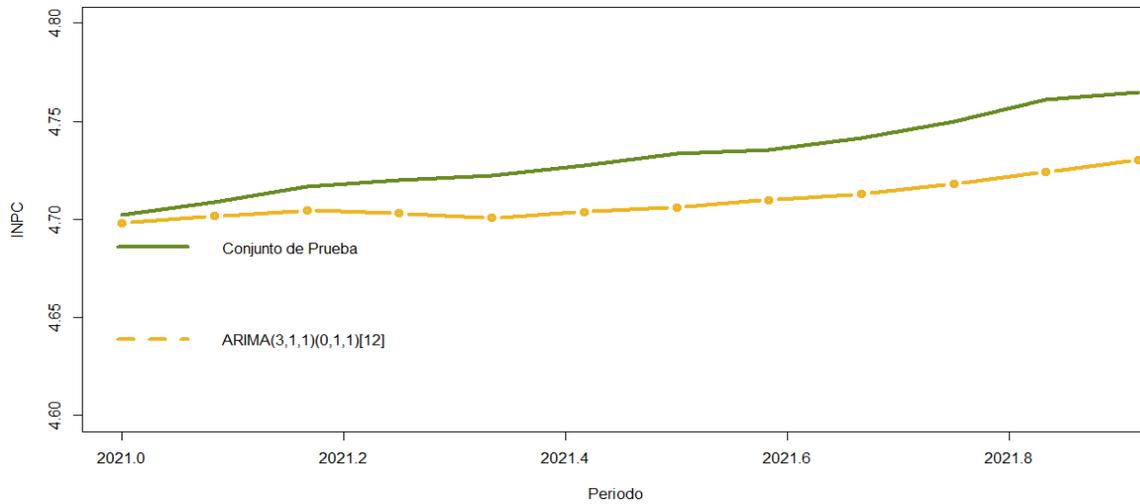
Fuente: Elaboración propia.

Gráfico 4.51. Comparación del conjunto de prueba vs modelo $SARIMA(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



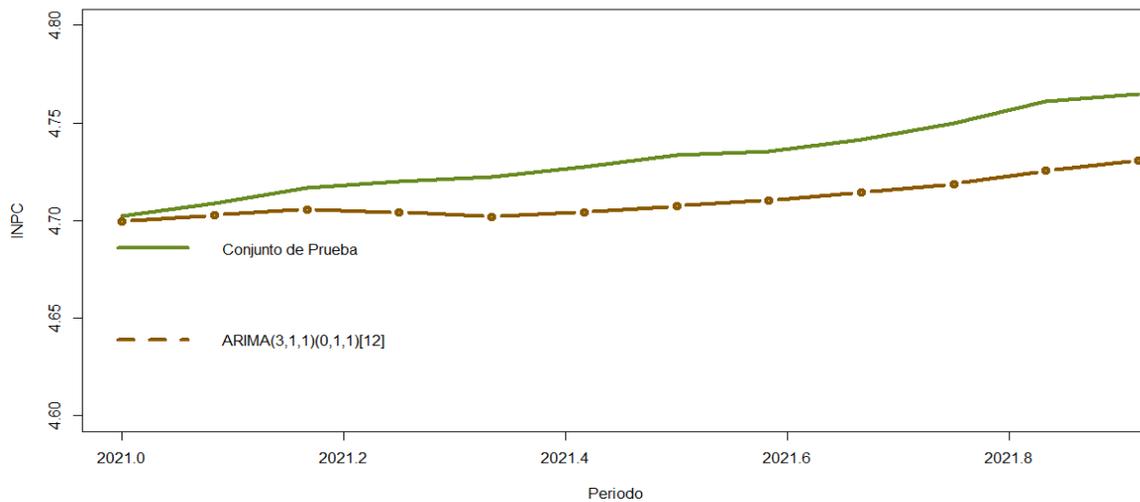
Fuente: Elaboración propia.

Gráfico 4.52. Comparación del conjunto de prueba vs modelo $SARIMA(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.



Fuente: Elaboración propia.

Gráfico 4.53. Comparación del conjunto de prueba vs modelo $SARIMA(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo y oferta y demanda agregada.



Fuente: Elaboración propia.

Tabla 4.11. Comparación de las medidas de error MAE, RMSE y MAPE y otros criterios, entre las ecuaciones planteadas con diferentes variables exógenas construidas con el conjunto de entrenamiento y prueba.

Ecuación	MAE	MAPE	RMSE	AIC
SARIMAX (3,1,1) (1,1,1) con las variables salario mínimo oferta y demanda agregada y expectativas inflacionarias	0.021	0.441	0.023	-2216.250
SARIMAX (4,1,3) (0,1,1) con las variables salario mínimo oferta y demanda agregada y expectativas inflacionarias	0.018	0.386	0.021	-2325.990
SARIMAX (3,1,1) (0,1,1) con las variables salario mínimo oferta y demanda agregada y expectativas inflacionarias	0.022	0.475	0.025	-2214.330
SARIMAX (3,1,1) (0,1,1) con las variables salario mínimo y oferta y demanda agregada.	0.022	0.455	0.024	-2193.030

Fuente: Elaboración propia.

Tabla 4.12. Coeficientes y significancia del modelo $SARIMAX(3, 1, 1)(1, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.

Coeficiente	Valor	Estadístico t	Valor p
AR1	1.229	12.0100	0.000
AR2	-0.456	-4.364	0.000
AR3	0.177	2.596	0.051
MA1	-0.850	-10.643	0.000

SAR1	-0.172	-2.037	0.043
SMA1	-0.761	-12.415	0.000
SALARIO	-0.025	-2.979	0.003
OFERTA Y DEMANDA	0.029	4.664	0.000
EXPECTATIVAS	-0.001	-4.629	0.000

Fuente: Elaboración propia.

Tabla 4.13. Coeficientes y significancia del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.

Coeficiente	Valor	Estadístico t	Valor p
AR1	1.192	17.609	0.000
AR2	0.335	3.664	0.000
AR3	-1.039	-11.901	0.000
AR4	0.305	4.627	0.000
MA1	-0.830	-16.447	0.000
MA2	-0.813	-10.683	0.000
MA3	0.994	19.435	0.000
SMA1	-0.803	-16.260	0.000
SALARIO	-0.024	-2.920	0.004
OFERTA Y DEMANDA	0.028	4.940	0.000
EXPECTATIVAS	-0.001	-4.342	0.000

Fuente: Elaboración propia.

Tabla 4.14. Coeficientes y significancia del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias.

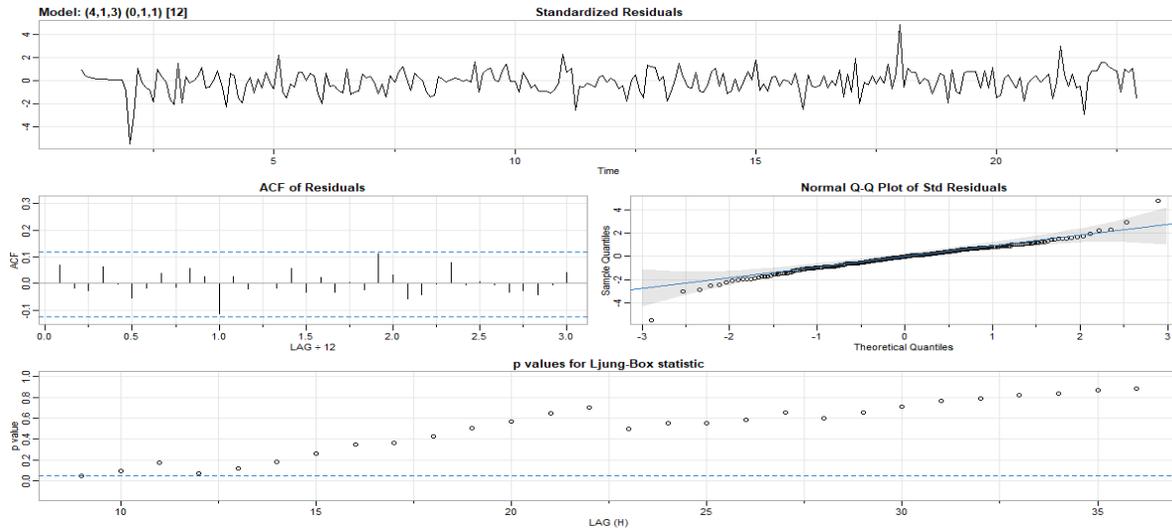
Coeficiente	Valor	Estadístico t	Valor p
AR1	1.244	10.781	0.000

AR2		-0.473	-4.441	0.000
AR3		0.176	2.578	0.011
MA1		-0.858	-8.927	0.000
SMA1		-0.828	-17.533	0.000
SALARIO		-0.021	-2.527	0.012
OFERTA	Y	0.030	4.797	0.000
DEMANDA				
EXPECTATIVAS		-0.001	-4.646	0.000

Fuente: Elaboración propia.

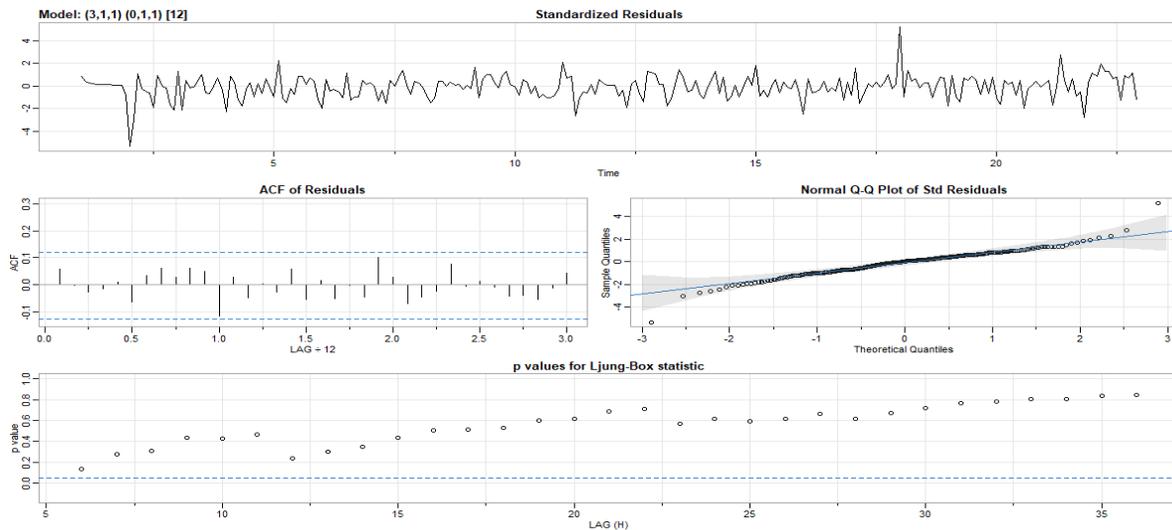
De acuerdo con el análisis comparativo el mejor modelo para realizar los pronósticos del año 2022 es el $SARIMAX(4,1,3)(0,1,1)_{12}$ con las variables exógenas salario mínimo, oferta y demanda agregada y expectativas inflacionarias. Sin embargo, es la ecuación con mayor número de parámetros. El segundo modelo con las medidas de error más bajas, así como el criterio de Akaike menor, es el $SARIMAX(3,1,1)(0,1,1)_{12}$ con las mismas variables exógenas y dos parámetros menos. Todas las ecuaciones presentadas cumplen con el requerimiento de significancia en sus parámetros, para estas dos ecuaciones el valor p aumenta muy poco para la variable salario en la segunda opción. Debido a los resultados estadísticos se tomarán en cuenta ambos modelos para realizar el pronóstico a un año, los resultados de los residuales de ambos modelos tomando en cuenta todo el conjunto de datos se muestran en los Gráficos 4.54 y 4.55.

Gráfico 4.54. Residuales del modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias tomando todo el conjunto de datos disponible.



Fuente: Elaboración propia.

Gráfico 4.55. Residuales del modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$ con las variables salario mínimo, oferta y demanda agregada y expectativas inflacionarias tomando todo el conjunto de datos disponible.



Fuente: Elaboración propia.

Al igual que los modelos realizados con el conjunto de entrenamiento, los residuales de este modelo siguen una distribución normal. Los rezagos se mantienen dentro de la banda

de confianza, para el primer modelo hay dos rezagos significativos, y uno de ellos se elimina con la opción de modelo número 2. En ambos casos se alcanza la independencia con valores p en la prueba Ljung Box de 0.5505 y 0.6151 respectivamente. Con los resultados de estos modelos se realizan los pronósticos para el 2022 del INPC con la transformación logarítmica y sin la transformación para conocer el valor real de la inflación. Los resultados se presentan en la Tabla 4.15.

Tabla 4.15. Pronósticos del INPC para el año 2022 con el modelo $SARIMAX(4, 1, 3)(0, 1, 1)_{12}$

Mes	Pronóstico con escala logarítmica	Pronóstico sin escala
Enero	4.775	118.523
Febrero	4.780	119.131
Marzo	4.785	119.726
Abril	4.785	119.735
Mayo	4.785	119.673
Junio	4.789	120.143
Julio	4.792	120.602
Agosto	4.796	121.038
Septiembre	4.801	121.593
Octubre	4.806	122.213
Noviembre	4.814	123.162
Diciembre	4.819	123.819

Fuente: Elaboración propia.

Tabla 4.16. Pronósticos para el año 2022 con el modelo $SARIMAX(3, 1, 1)(0, 1, 1)_{12}$

Mes	Pronóstico con escala logarítmica	Pronóstico sin escala
Enero	4.776	118.606
Febrero	4.780	119.138
Marzo	4.785	119.739
Abril	4.785	119.675
Mayo	4.784	119.592
Junio	4.788	120.006
Julio	4.792	120.484
Agosto	4.795	120.925
Septiembre	4.800	121.542
Octubre	4.805	122.163
Noviembre	4.814	123.167
Diciembre	4.819	123.850

Fuente: Elaboración propia.

Con los pronósticos del INPC realizados se procede a obtener la variación mensual anualizada de la inflación respecto al año 2021 con la fórmula (2) descrita y utilizada para el modelo SARIMA. Los resultados se muestran en la Tabla 4.17.

Tabla 4.17. Pronósticos de la inflación mensual anualizada en México con las variables exógenas salario mínimo, oferta y demanda agregada y expectativas inflacionarias.

Mes	$SARIMAX(4, 1, 3)(0, 1, 1)_{12}$	$SARIMAX(3, 1, 1)(0, 1, 1)_{12}$
Enero	7.54%	7.62%
Febrero	7.41%	7.42%
Marzo	7.07%	7.08%
Abril	6.72%	6.67%
Mayo	6.45%	6.38%
Junio	6.30%	6.18%
Julio	6.09%	5.98%
Agosto	6.27%	6.17%
Septiembre	6.10%	6.06%
Octubre	5.76%	5.71%
Noviembre	5.37%	5.38%
Diciembre	5.55%	5.58%

Fuente: Elaboración propia.

Los pronósticos con ambas ecuaciones cumplen con los objetivos de BANXICO. El segundo modelo empieza con valores más altos para los primeros tres meses del año y después su valor queda un poco por debajo de los pronosticados con la primera ecuación, ambos modelos utilizan las mismas variables exógenas, por lo que la diferencia se encuentra en el número de parámetros de cada ecuación. En modelos de series de tiempo por el principio de parsimonia se recomienda la ecuación que contenga menos, ambos modelos tienen buenos resultados y pronósticos y son próximos al valor real de la inflación en los primeros meses, a partir de agosto los valores proyectados resultaron más alejados.

4.3 Modelo de Red Neuronal

Tomando en cuenta lo mencionado en el capítulo 3.3.4 se cuenta con las herramientas suficientes para poder elaborar nuestros modelos de predicción de perceptrones multicapa para poder predecir el INPC, para lo cual se elaboraron modelos a través de Python en los cuales tenemos las siguientes variables:

Y = INPC al tiempo t+1.

X1 = INPC al tiempo t.

X2 = Expectativas al tiempo t (Se presenta como una variable dicotómica).

X3 – X14 = Variables dummy sobre los meses (Se presenta como una variable dicotómica).

Una vez que se tiene la matriz de entrada y el vector de salida que sería “Y”, pasaríamos a transformar los datos del INPC de la siguiente manera para poder normalizarlos justo como lo hace Héctor Pedroza en su artículo “Uso de redes neuronales para el pronóstico de la inflación en México”:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

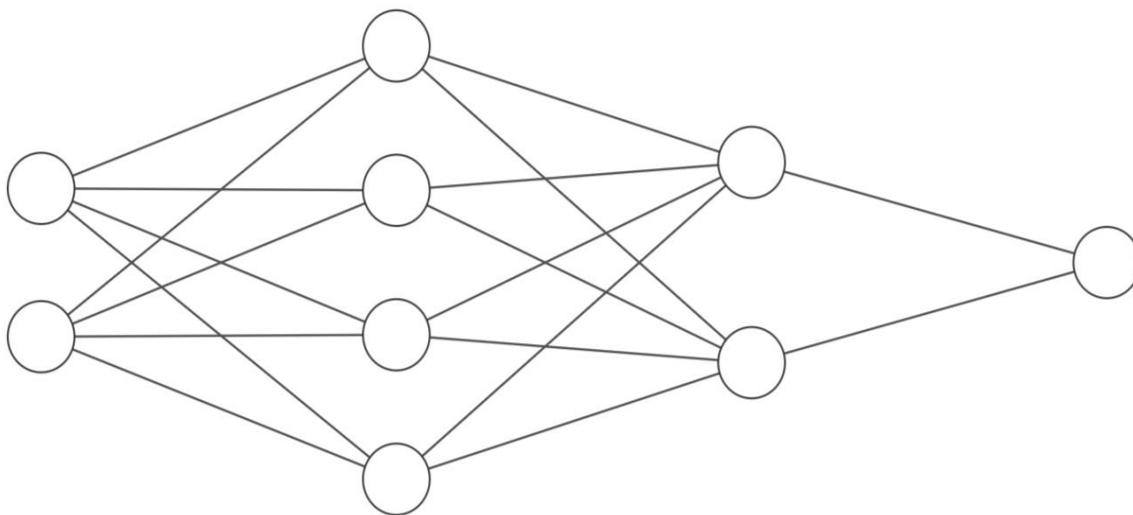
Una vez con los datos escalados para “X1” y “Y” puede iniciarse el proceso de construcción de los modelos de perceptrón múltiples.

4.3.1. Perceptrón múltiple utilizando las expectativas y el INPC al tiempo t.

Para este modelo se ocupó un perceptrón multicapa con las siguientes características:

- Capa de entrada: 2 neuronas.
- Primera capa oculta: 4 neuronas.
- Segunda capa oculta: 2 neuronas.
- Capa de salida: 1 neurona.

Gráfico 4.56. Perceptrón multicapa conformado por dos neuronas de entrada, dos capas ocultas de 4 y 3 neuronas con una neurona en la capa de salida.



Fuente: Elaboración propia a través del aplicativo web NN-SVG.

Para la elaboración de los perceptrones multicapa se utilizó el software de Google Colab que es basado en Python, esto con la finalidad de poder usar las paqueterías de tensorflow y keras las cuales nos ayudaran a elaborar los modelos correspondientes.

Al momento de armar el modelo es importante establecer un learning ratio el cual nos ayudara a dar un parámetro de ajuste cada vez que el modelo requiera realizar ajustes. Tomando en cuenta que entre más pequeño sea más tardara la red neuronal en terminar de aprender y ajustar sobre sí misma y necesitara más épocas²² de aprendizaje por lo que

²² Número de iteraciones necesarias para que la red neuronal aprenda y minimice el error cuadrado medio.

para este modelo se establece un $lr = 0.001$. Así mismo estableceremos el número de épocas que se ejecutaran para que nuestro perceptrón pueda aprender de sí mismo a través del algoritmo ya mencionado de retro propagación, estableciendo que se usaran un total de 8000 épocas para este modelo.

Una vez que se tienen estos parámetros establecidos se pasara a construir el modelo dentro del código, una vez establecido se puede empezar a entrenar y observar los resultados. Para poder entrenar una red neuronal con estas características es importante separar a los datos en el set de entrenamiento y el set de pruebas o de testeo, con la intención de corroborar que la red neuronal ha aprendido correctamente y no ha habido un sobrentrenamiento²³, para lo que se usara un total del 90% de los datos para formar el set de entrenamiento dejando el año de 2021 como el set de prueba con la intención de que la red neuronal ha aprendido correctamente.

Para este modelo las variables de entrada son las expectativas y el INPC al tiempo t , tomando en cuenta que las expectativas están entendidas como una variable dicotómica en donde:

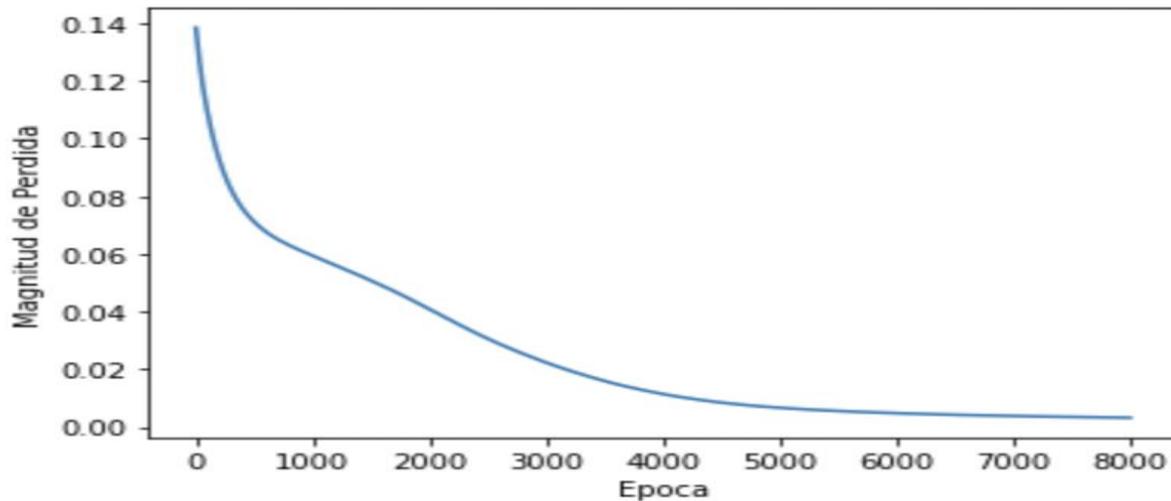
1 = La expectativa fue mas alta que el valor real de la inflación

0 = La expectativa fue menor al valor real de la inflación

Una vez establecidos los parámetros, se aplicará el modelo con ayuda de las paqueterías de tensorflow y keras mencionadas anteriormente, con 8,000 épocas de aprendizaje el código arroja el siguiente gráfico.

²³ En lugar de que la red aprenda sobre sí misma, memoriza los datos y los replica.

Gráfico 4.57. Grafica de magnitud de pérdida en el entrenamiento.



Fuente: Elaboración propia a través de Google Colab.

Como puede apreciarse en el Gráfico 4.57 el haber seleccionado un total de 8,000 épocas ayuda para que al modelo le dé tiempo suficiente de aprender sobre sí mismo llegando a una pérdida de casi 0, lo que nos ayuda a ver que con estas épocas es más que suficiente para su aprendizaje.

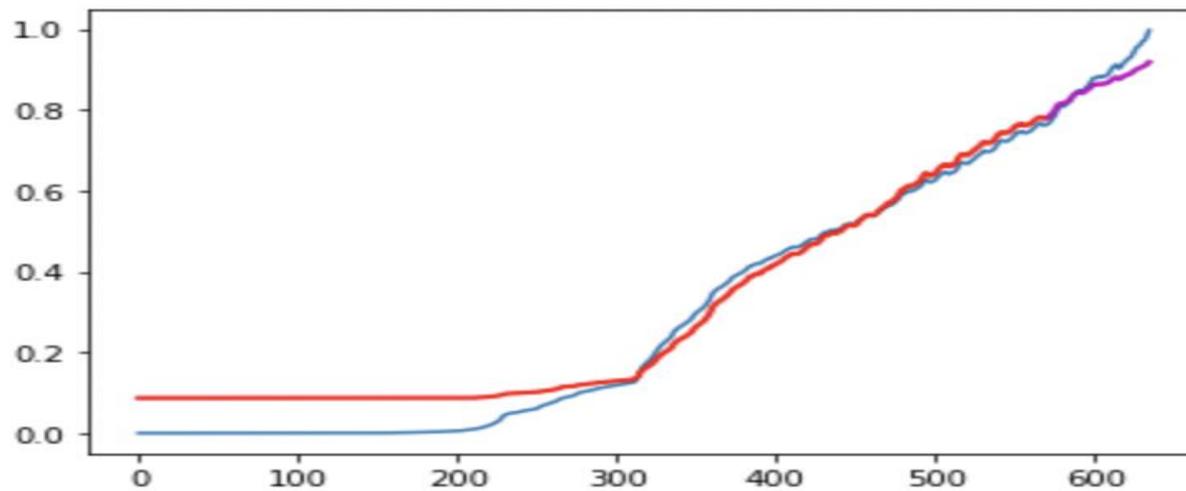
Ahora lo que se tendrá que analizar es la precisión del modelo (*acc*), así como la pérdida (*loss*) de acuerdo con los siguientes resultados del modelo:

$$\mathbf{loss = 0.00114180985}$$

$$\mathbf{acc = 0.01562}$$

Como se observa la precisión del modelo es muy baja, apenas 1.5% sobre los datos del conjunto de prueba, aunque si hay que tomar algo en cuenta es que la pérdida es muy baja por lo que es bastante buena para este modelo, aunque no sea preciso sobre los datos no está tan lejos de ellos. Tomando todo esto en cuenta se realiza un análisis gráfico sobre el comportamiento de los ajustes al modelo (Gráfico 4.58).

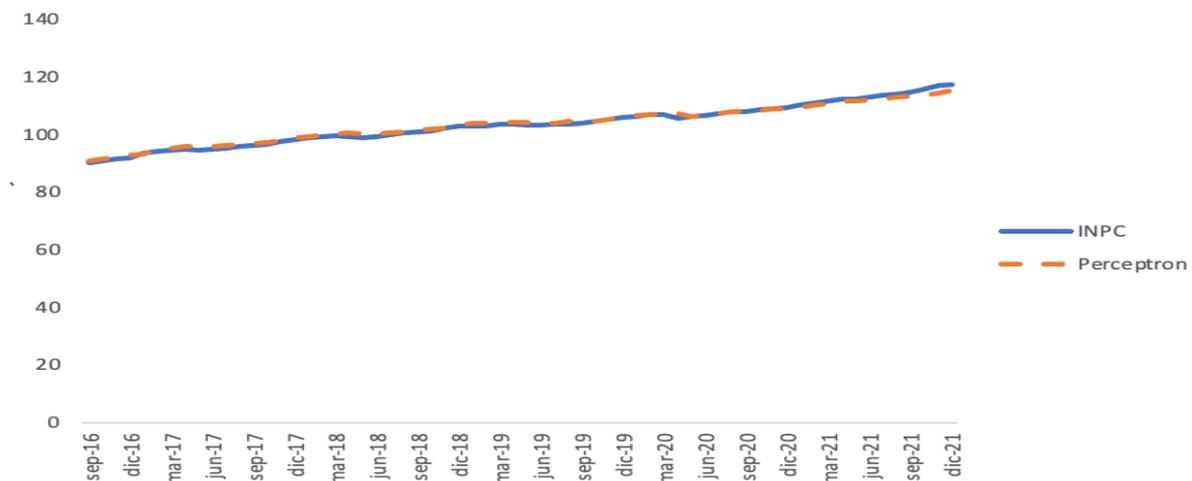
Gráfico 4.58. Grafica de comparación entre datos originales, entrenamiento y prueba.



Fuente: Elaboración propia a través de Google Colab

El ajuste es bastante bueno a nivel gráfico y tomando en cuenta el hecho de que, aunque la precisión del modelo fue baja, la diferencia entre los datos es mínima, y a su vez puede verse que los datos de prueba se acercan bastante como se aprecia en el Gráfico 4.59.

Gráfico 4.59. Comparativa: INPC y datos simulados por el perceptrón.



Fuente: Elaboración propia.

En conclusión, la predicción es bastante buena, aunque al final el modelo se sigue quedando algo corto ya que el INPC real empieza a despegar y a alejarse de los datos predichos con los datos de prueba (Tabla 4.18).

Tabla 4.18. Predicción del perceptrón vs el INPC real

Fecha	01/07/21	01/08/21	01/09/21	01/10/21	01/11/21	01/12/21
INPC	113.682	113.899	114.601	115.561	116.884	117.308
Predicción	106.125	106.404	106.494	106.779	107.159	107.662

Fuente: Elaboración propia con los outputs del modelo de red neuronal perceptrón (2,4,2,1).

Así, en el último año de predicción los datos se comienzan a alejar bastante ya que el INPC aumento de manera considerable en este periodo de tiempo.

4.3.2 Modelo con expectativas y dummies sobre los meses del año.

Para este segundo modelo se agregaron las variables dummy para los meses, con la finalidad de modelar la estacionalidad del INPC. De forma simple se agregaron estas variables a la matriz de entrada de la siguiente forma (tablas 4.19 y 4.20).

Tabla 4.19. Tabla de insumo que se utiliza para la estacionalidad de los meses

Enero	Febrero	Marzo	Abril	Mayo	Junio
1	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Fuente: Elaboración propia en Excel 365.

Tabla 4.20. Tabla de insumo que se utiliza para la estacionalidad de los meses

Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1

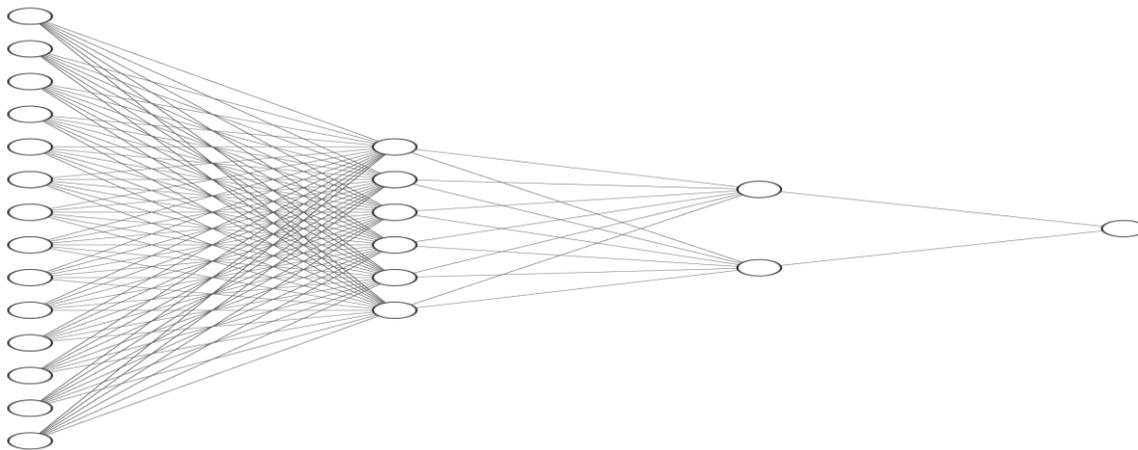
Fuente: Elaboración propia en Excel 365.

Cabe destacar que esto se agregó para todos los años correspondientes en el periodo que estamos abarcando, dejando una matriz con un total de 14 entradas que se puede interpretar como un modelo con una capa de entrada conformada por 14 neuronas.

Para este segundo modelo se utilizará un perceptrón multicapa el cual tendrá la siguiente estructura (gráfico 4.60):

- Capa de entrada: 14 neuronas.
- Primera capa oculta: 6 neuronas.
- Segunda capa oculta: 2 neuronas.
- Capa de salida: 1 neurona.

Gráfico 4.60. Perceptrón (14,6,2,1).



Fuente: Elaboración propia a través del aplicativo web NN-SVG.

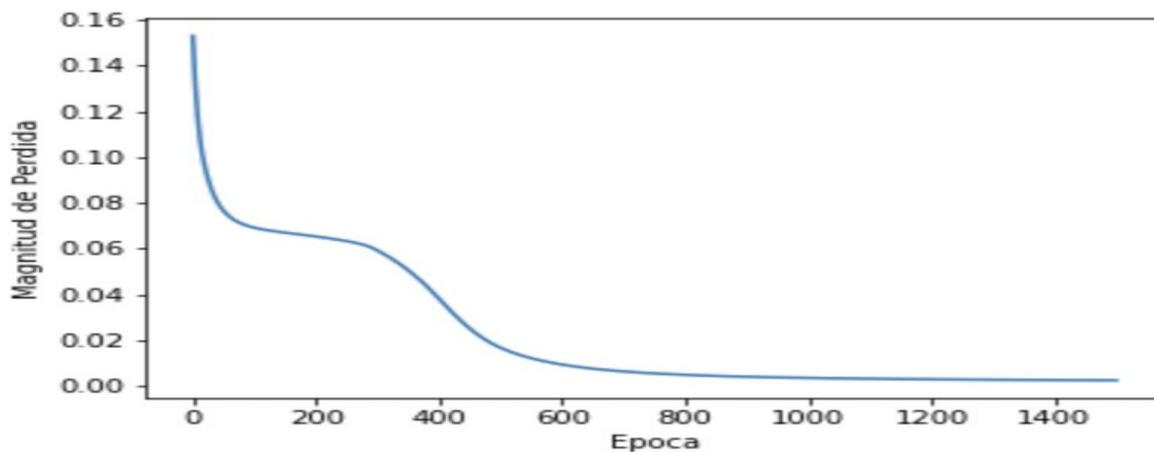
Para este modelo se han ajustado los parámetros sobre el entrenamiento, dejándolos de la siguiente manera:

$$lr = 0.01$$

$$epochs = 1500$$

Al estar haciendo los ajustes sobre este modelo se ha notado que aumentando el valor sobre el learning ratio podemos disminuir el número de épocas de entrenamiento para que la pérdida comience a aproximarse a cero, dando como resultado la siguiente el Gráfica 4.61 de magnitud.

Gráfico 4.61. Magnitud de pérdida en el entrenamiento.



Fuente: Elaboración propia a través de Google Colab.

Como puede observarse la convergencia es bastante buena en un número significativamente reducido de épocas en comparación con el modelo anterior. Derivado de lo anterior, este modelo proporciona los siguientes parámetros en cuanto a su pérdida y precisión:

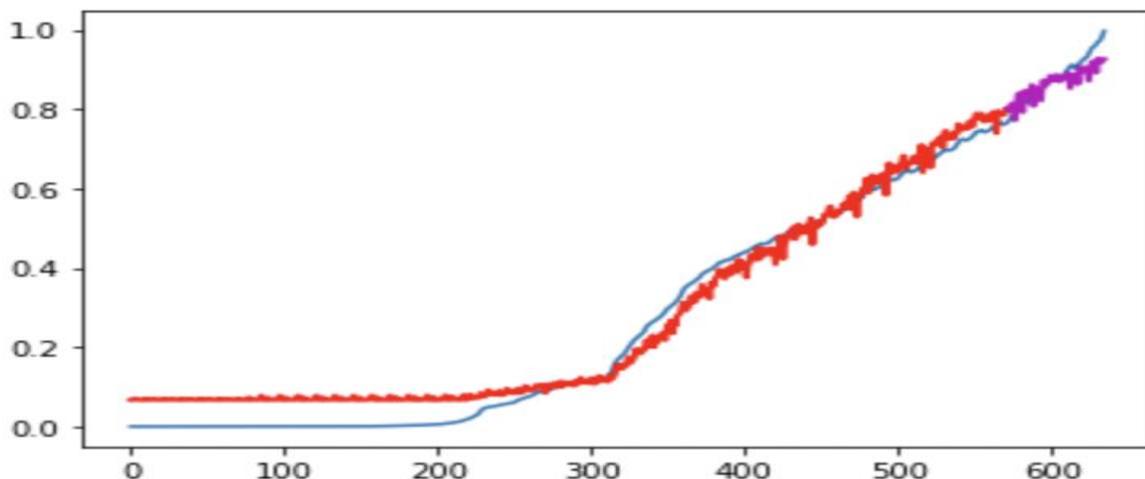
$$loss = 0.001046841032$$

$$acc = 0.015625$$

Así, la pérdida es incluso más pequeña comparada con el modelo anterior, aunque la precisión del modelo sigue siendo prácticamente la misma. Analizando estos resultados puede deducirse que este modelo tiene un mejor comportamiento que el anterior.

Como era de esperarse y a simple vista este modelo es aún más complejo que el anterior por el simple hecho de tener más neuronas que componen al modelo. A simple vista y con todo lo que se ha observado sobre este modelo parece ser bastante bueno e incluso mejor que el primero, aunque todavía habrá que evaluar su poder de predicción y ajuste gráfico por lo que se procede a elaborar la comparación en el Gráfico 4.62.

Gráfico 4.62. Comparación entre datos originales, entrenamiento y prueba.

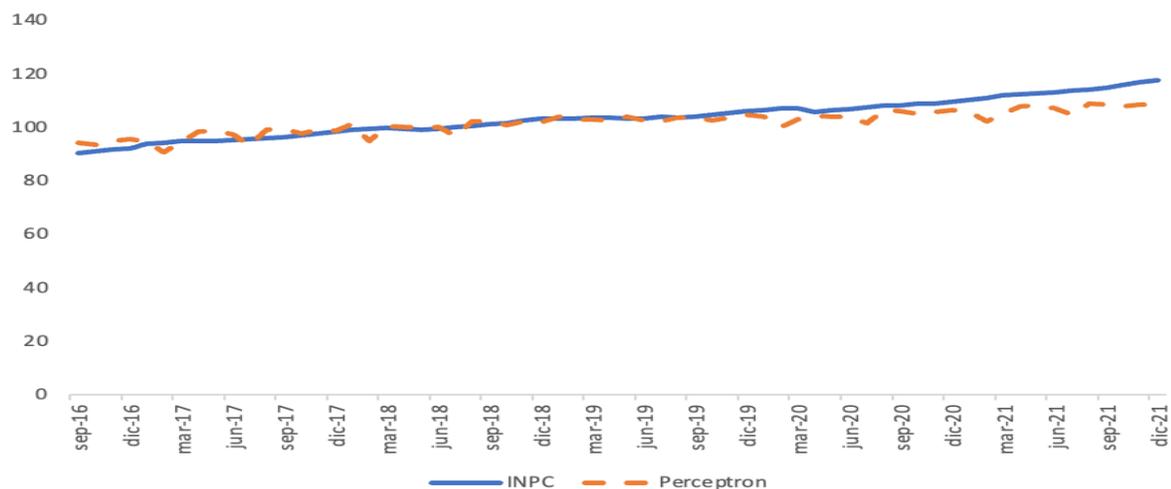


Fuente: Elaboración propia a través de Google Colab.

Podemos visualizar que el ajuste es bastante bueno incluso con los datos del conjunto prueba, pero así mismo se puede visualizar que existen variaciones bastante significativas en el ajuste de una manera negativa lo que nos indica que al momento de predecir o

ajustar, está dándonos valores negativos sobre el INPC lo que suele ser raro en esta variable ya que si existen variaciones negativas no son tan comunes. Dado el gráfico anterior, puede verse de una manera más precisa el periodo de prueba para corroborar las variaciones que existen en el modelo.

Gráfico 4.63. Comparativa: INPC y datos simulados por el perceptrón.



Fuente: Elaboración propia.

Como podemos observar en el **Gráfico 4.63** tenemos variaciones negativas y positivas importantes que pueden afectar drásticamente al modelo, haciéndolo demasiado volátil siendo el INPC un indicador que no suele tener variaciones tan significantes de manera negativa.

Tabla 4.21. Predicción del perceptrón vs el INPC real

Fecha	01/07/21	01/08/21	01/09/21	01/10/21	01/11/21	01/12/21
INPC	113.682	113.899	114.601	115.561	116.884	117.308
Predicción	104.619	108.751	108.370	107.525	108.489	108.561

Fuente: Elaboración propia con los outputs del modelo de red neuronal perceptrón (14,6,2,1).

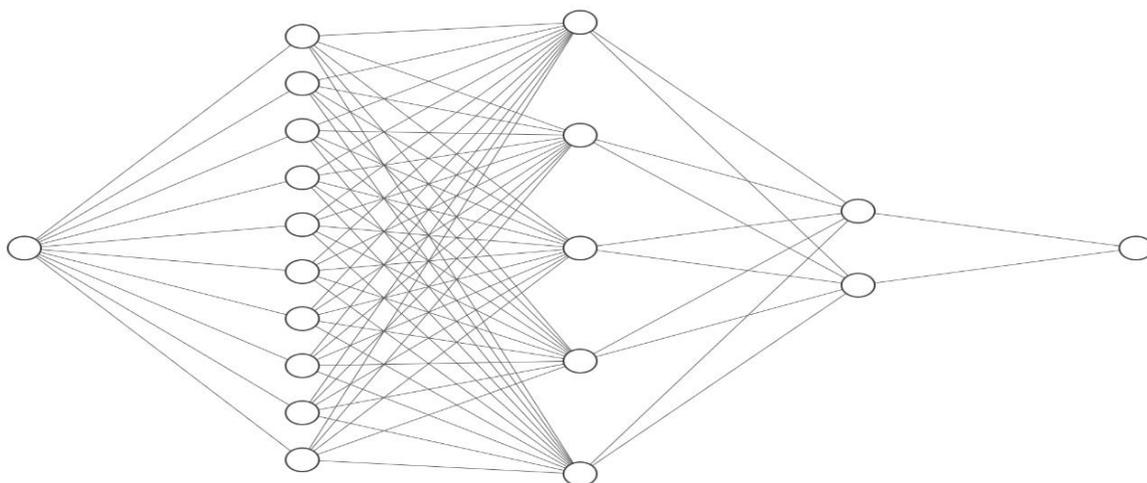
Según los resultados de la tabla anterior, sigue quedándose muy por debajo del INPC real aparte de que tiene variaciones importantes como se puede observar del mes de julio a octubre, por lo que en manera de conclusión para este modelo es que requiere menos épocas de entrenamiento y con un learning ratio más grande por lo que el tiempo de cómputo es bastante optimo, aunque las variaciones negativas que tiene no cuadran con el comportamiento real que puede tener el INPC.

4.3.3. Modelo utilizando únicamente el INPC al tiempo t

Este tercer modelo es todavía más simple que los modelos anteriores ya que nos basaremos únicamente en la información del INPC para predecirlo, como un modelo de series de tiempo, solo que en esta ocasión usaremos las redes neuronales, por lo que la estructura del modelo en cuanto a sus capas será la siguiente:

- Capa de entrada: 1 neuronas.
- Primera capa oculta: 10 neuronas.
- Segunda capa oculta: 5 neuronas.
- Tercera capa oculta: 2 neuronas.
- Capa de salida: 1 neurona.

Gráfico 4.64. Perceptrón (1,10,5,2,1)



Fuente: Elaboración propia a través del aplicativo web NN-SVG.

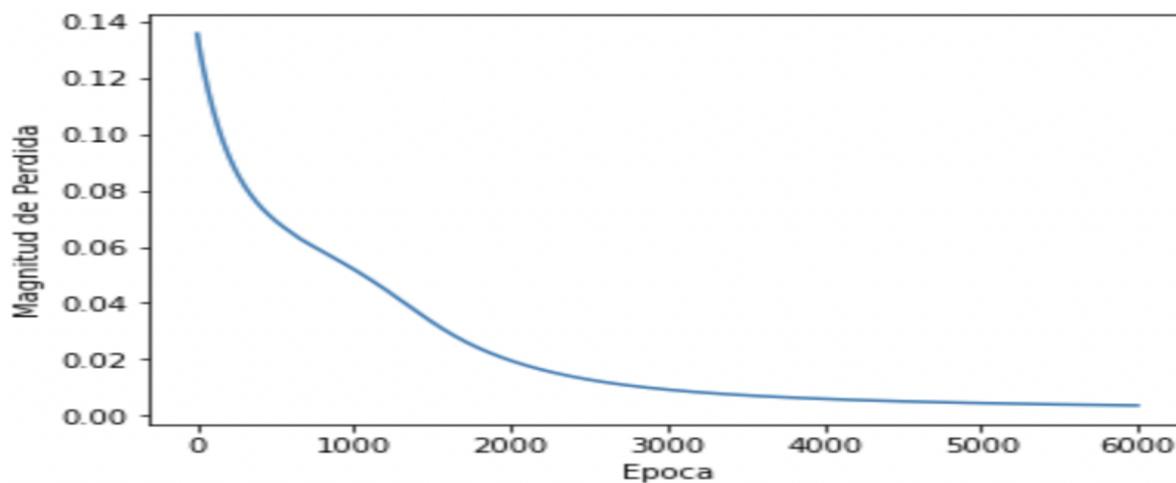
Los parámetros que se utilizarán para este modelo serán los siguientes, los cuales son muy similares a los del primer modelo:

$$lr = 0.001$$

$$epochs = 6000$$

Como se observa, pese a que solo se ocupará una sola neurona de entrada, se han agregado varias dentro de las capas ocultas para poder tener un mejor ajuste, incluso se ocupó un learning ratio igual al del primer modelo.

Gráfico 4.65. Magnitud de pérdida en el entrenamiento.



Fuente: Elaboración propia través de Google Colab.

Podemos observar incluso que la curva en la gráfica de magnitud de pérdida es todavía más suave, y llega a ajustar lo suficiente con las 6000 épocas. Al momento de entrenar esta red neuronal nos ha arrojado los siguientes parámetros en cuanto a precisión y pérdida:

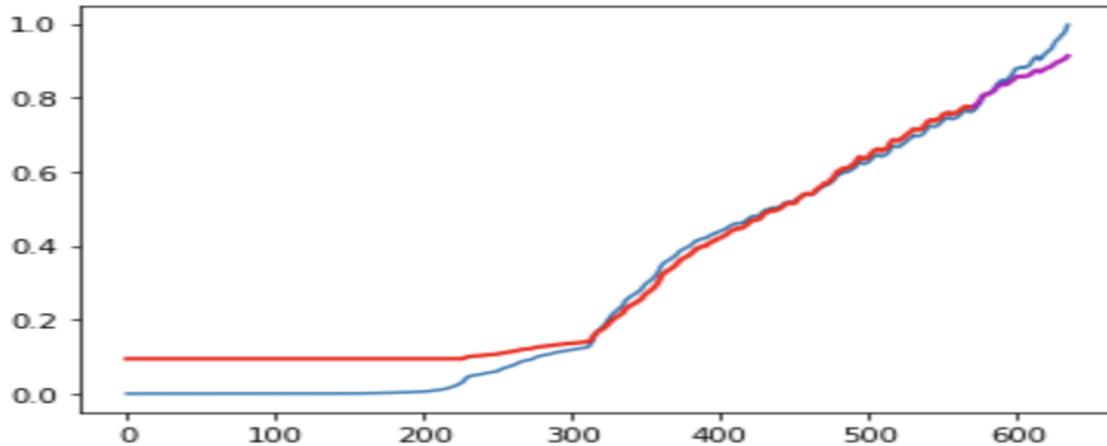
$$loss = 0.0014568474143$$

$$acc = 0.015625$$

Podemos observar que la precisión se ha mantenido igual para los tres modelos, y lo que ha cambiado de alguna manera significativa ha sido la pérdida, que, en comparación con los anteriores dos modelos, podríamos decir que este modelo tiene una pérdida mayor a los anteriores.

Analizando el comportamiento que ha tenido, nos da una primera impresión de que puede ser un modelo con menor eficiencia en comparación con los anteriores por lo que se procederá a analizar la gráfica en la que se observan los datos reales contra los datos entrenados y de prueba.

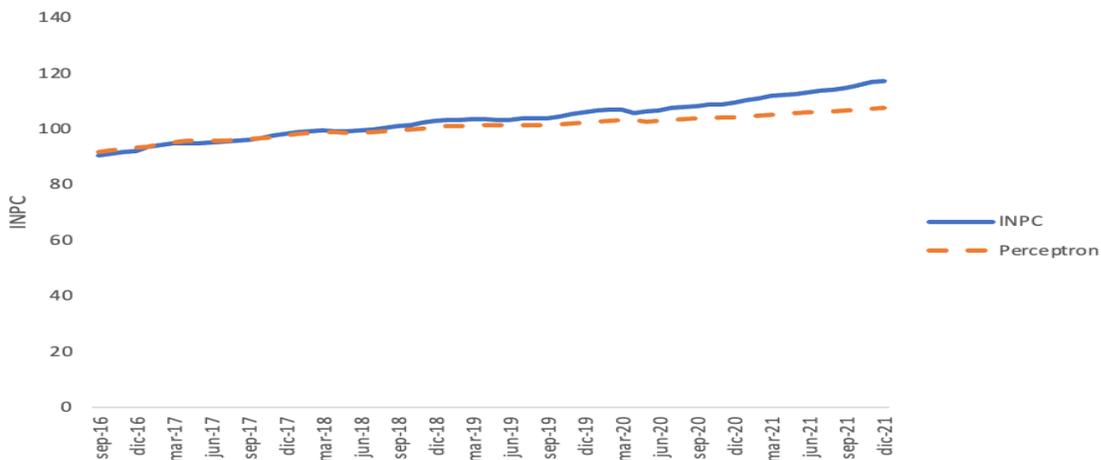
Gráfico 4.66. Comparación entre datos originales, entrenamiento y prueba.



Fuente: Elaboración propia a través de Google Colab.

Podemos observar que es muy parecido al primer modelo y no tiene el problema del segundo con las variaciones negativas, pero si se puede observar que en la parte donde se hace la comparación entre el periodo de prueba y el INPC esta aún más despegado que el primer modelo.

Gráfico 4.67. Comparativa: INPC y datos simulados por el perceptrón.



Fuente: Elaboración propia a través de Excel 365.

De lo anterior, se observa que las variaciones que tiene son menos en comparación al segundo modelo, pero al igual que los dos modelos anteriores queda rezagado en los últimos meses del año 2021, lo que sigue siendo un problema ya que conforme el INPC real comienza a aumentar en el modelo se va quedando rezagado como se puede observar en la Tabla 4.22.

Tabla 4.22. Predicción del perceptrón vs el INPC real

Fecha	01/07/21	01/08/21	01/09/21	01/10/21	01/11/21	01/12/21
INPC	113.682	113.899	114.601	115.561	116.884	117.308
Predicción	105.437	105.722	105.813	106.105	106.493	107.009

Fuente: Elaboración propia con los outputs del modelo de red neuronal perceptrón (1,10,5,2,1).

En la tabla anterior se observa que los datos se van separando conforme aumenta el INPC en los últimos meses del 2021, dejando que el modelo se quede bastante rezagado de lo que es la realidad, por lo que la manera de concluir este tercer modelo es diciendo que, pese a que se queda varias unidades por debajo de lo esperado tiene la ventaja sobre el segundo modelo de no tener las variaciones negativas tan grandes que tiene, por lo que podríamos catalogarlo como el segundo mejor modelo de los tres aunque la diferencia con el primero no es demasiada.

Analizando los tres modelos se han podido observar diferentes comportamientos al ajustar los parámetros del perceptrón por lo que se ha llegado a la conclusión de que al tomar un learning ratio más grande el tiempo de cálculo es menor y los modelos requieren menos épocas de aprendizaje como se apreció en el tercer modelo. De la misma manera, puede analizarse que estos se quedan por debajo del valor esperado para 2021 ya que se observa que no ajusta lo suficiente al final del periodo por tratarse de aumentos significativos en el INPC.

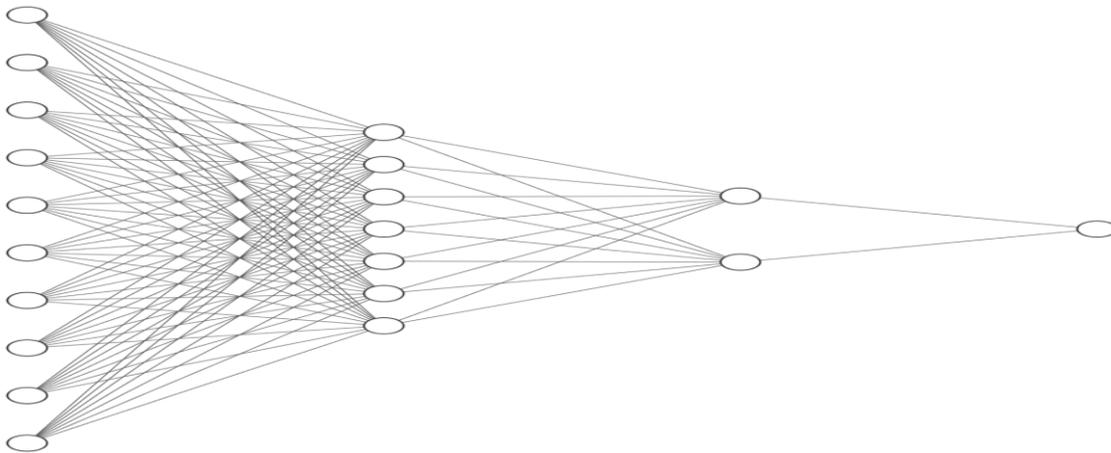
4.3.4. Modelo perceptrón ajustado en base a la experiencia de los modelos anteriores

Con base a la experiencia de los modelos anteriores se decidieron hacer varios ajustes al modelo siendo uno de los principales el cómo se escalaron los datos ya que se notó que las predicciones no llegaban más allá del INPC más alto reportado ya que este era el límite, por lo que se decidió tomar un INPC máximo de 150 para poder escalar el modelo y dar apertura a que existan valores más allá del INPC máximo reportado.

Tomando esto en cuenta se estableció un perceptrón con la siguiente estructura:

- Capa de entrada: 1 neuronas.
- Primera capa oculta: 10 neuronas.
- Segunda capa oculta: 7 neuronas.
- Tercera capa oculta: 2
- Capa de salida: 1 neurona

Gráfico 4.68. Perceptrón (1,10,7,2,1).



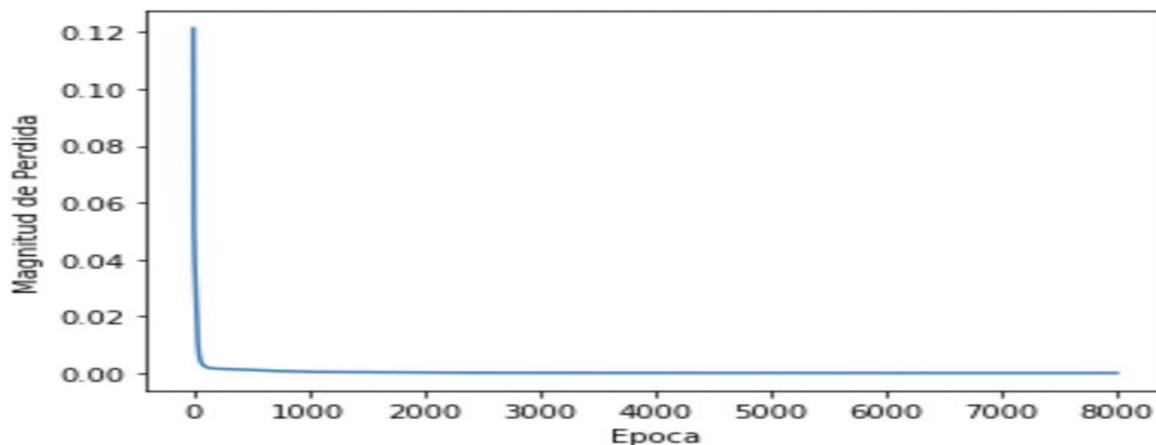
Fuente: Elaboración propia a través del aplicativo web NN-SVG.

Los parámetros que se utilizarán para este modelo serán los siguientes, los cuales son muy similares a los del primer modelo:

$$lr = 0.1$$
$$epochs = 8000$$

Como ya se había mencionado al tener un learning ratio mayor colabora a que el ajuste sea aún más rápido y le cueste mucho menos a la red neuronal entrenarse lo que nos lleva a mejores resultados y un mejor ajuste aún más rápido por lo que también el tener 8000 épocas de entrenamiento se podrá a ajustar el modelo de una forma más precisa (gráfico 4.69).

Gráfico 4.69. Magnitud de pérdida en el entrenamiento



Fuente: Elaboración propia través de Google Colab.

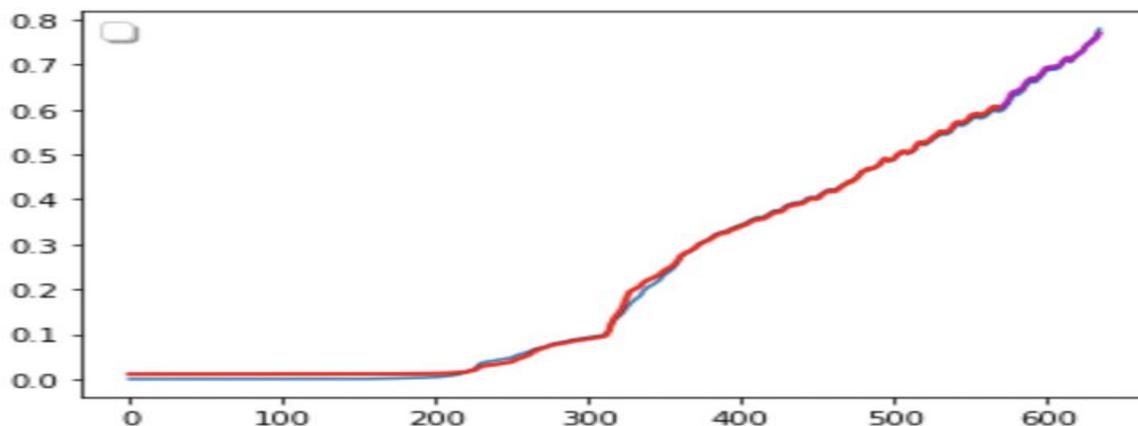
Como puede confirmarse, el ajuste fue bastante rápido en un numero de épocas menor a 1000 y solo siguió disminuyendo a niveles de centésimas con el número de épocas que se siguió entrenando el modelo, pero realmente no aprendió mucho más. En cuanto a precisión y perdida se definieron los siguientes parámetros que nos arrojó el modelo aplicado a los datos de testeo:

$$loss = 3.131409175693989e - 05$$

$$acc = 0$$

Que, aunque la precisión es de cero se tiene una perdida casi igual a cero lo que indica que el modelo está bien ajustado y la diferencia en realidad es mínima. Teniendo en cuenta los datos que se obtuvieron con anterioridad se procede a hacer la comparación gráfica con los datos de entrenamiento y los de prueba para poder visualizar si el entrenamiento fue correcto como a su vez verificar que no se trata de un modelo sobreentrenado (gráfico 4.70).

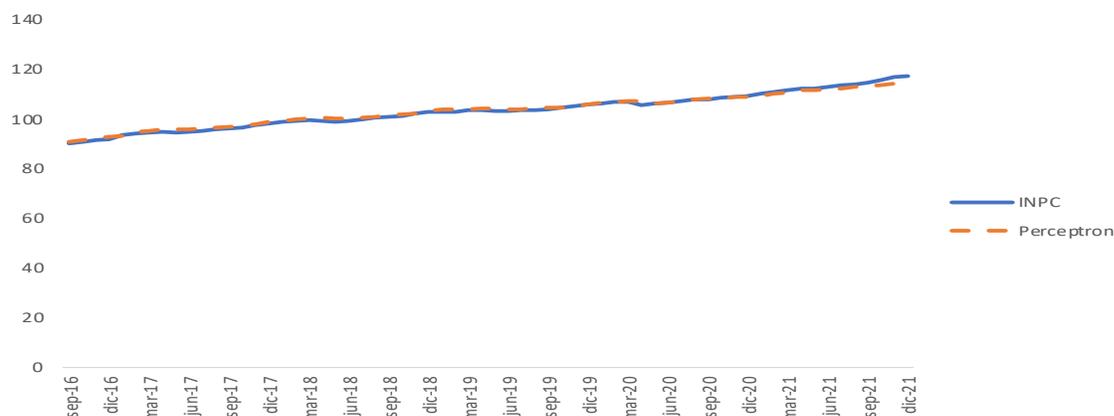
Gráfico 4.70. Comparación entre datos originales, entrenamiento y prueba.



Fuente: Elaboración propia a través de Google Colab.

Como se puede comprobar, los datos entrenados (línea roja) ajustan aún mucho mejor a los datos originales desde un principio a comparación de los modelos anteriores, así como también los datos de prueba (línea morada) de igual forma ajusta de una forma muy eficiente a los datos reales dando por sentado que el modelo aprendió y ajusto correctamente sin caer en un sobre entrenamiento. Tomando lo anterior en cuenta se visualizan los datos más de cerca ya que puede observarse que al final parece que se sigue quedando ligeramente por debajo del INPC real (gráfico 4.71).

Gráfico 4.71. Comparativa: INPC y datos simulados por el perceptrón.



Fuente: Elaboración propia a través de Excel.

Así, el modelo que mejor se ha ajustado hasta ahora dando unas predicciones bastantes precisas, aunque al final del periodo de diciembre del 2021 se ve que comienza a verse una diferencia significativa (Tabla 4.23).

Tabla 4.23. Predicción del perceptrón vs el INPC real

Fecha	01/07/21	01/08/21	01/09/21	01/10/21	01/11/21	01/12/21
INPC	113.682	113.899	114.601	115.561	116.884	117.308
Predicción	112.483	113.015	113.188	113.744	114.495	115.511

Fuente: Elaboración propia con los outputs del modelo de red neuronal perceptrón (1,10,7,2,1).

Puede observarse que conforme se acerca el 2022 el INPC del modelo comienza a quedarse por debajo de lo esperado, pero aun así este sigue siendo el modelo que más se acercó. Tomando en cuenta los resultados anteriores se procedió a calcular la tasa de inflación usando las predicciones del modelo para 2020 tomando como dato de inicio el INPC de enero de ese año, así como el antecedente de diciembre del 2019. Para calcular la inflación anual tomando en cuenta que las predicciones en esta ocasión son sobre los mismos datos de entrada que el modelo va arrojando.

Tabla 4.24. Predicción del perceptrón vs el INPC real utilizando la misma información de salida del perceptrón como entrada para el 2020.

Mes	INPC	Variación porcentual	Inflación Anual
01/12/19	105.934	NA	NA
01/01/20	106.447	0.484%	0.484%
02/01/20	106.942	0.465%	0.949%
03/01/20	107.376	0.406%	1.355%
04/01/20	107.755	0.353%	1.708%
05/01/20	108.083	0.305%	2.013%
06/01/20	108.367	0.262%	2.275%
07/01/20	108.610	0.225%	2.500%
08/01/20	108.819	0.192%	2.692%
09/01/20	108.998	0.164%	2.856%
10/01/20	109.150	0.140%	2.996%

11/01/20	109.280	0.119%	3.115%
12/01/20	109.483	0.186%	3.301%

Fuente: Elaboración propia en Excel con los datos de salida del perceptrón.

Los resultados indican que la inflación es muy cercana a la real ya que según datos del INEGI para diciembre de 2020 existió una inflación anual del 3.150%, por lo que a este modelo está prediciendo de forma muy acertada.

Tomando en cuenta lo anterior se predicen de la misma forma la inflación para 2022 tomando como dato de inicio el INPC de diciembre del 2021.

Tabla 4.25. Predicción del perceptrón vs el INPC real utilizando la misma información de salida del perceptrón como entrada para el 2022.

Mes	INPC	Variación porcentual	Inflación Anual
01/12/21	117.308	NA	NA
01/01/22	115.833	-1.2574%	-1.257%
02/01/22	114.705	-0.974%	-2.231%
03/01/22	113.826	-0.767%	-2.998%
04/01/22	113.130	-0.612%	-3.609%
05/01/22	112.573	-0.493%	-4.102%
06/01/22	112.122	-0.400%	-4.502%
07/01/22	111.756	-0.327%	-4.828%
08/01/22	111.456	-0.268%	-5.097%
09/01/22	111.210	-0.221%	-5.318%
10/01/22	111.006	-0.183%	-5.501%
11/01/22	110.838	-0.152%	-5.653%
12/01/22	110.698	-0.126%	-5.779%

Fuente: Elaboración propia en Excel con los datos de salida del perceptrón.

En este sentido, la inflación anual parece tener más el comportamiento de una deflación quedándose por muy lejos de la tasa de inflación anual del 8.7% a septiembre de 2022 que reporta el INEGI un comportamiento que se veía acercarse al analizar las gráficas y datos predichos para 2021 tomando esos mismos de entrada que el INPC del modelo tiende a quedarse por debajo para los años venideros después de 2020.

4.3.5. Red Neuronal Autorregresiva

Para el caso de la red neuronal autorregresiva se procede de la misma manera para ajustar el modelo pertinente, es decir, se divide el conjunto de datos disponible respetando la metodología que se llevó a cabo con el conjunto en la sección 4.1, conservando como datos de entrenamiento las observaciones del Índice Nacional de Precios al Consumidor de enero de 1969 hasta diciembre de 2020 dejando un total de 624 observaciones para entrenar a la red correspondiente y dejando 12 como conjunto de prueba para medir la eficiencia de pronóstico de la red ajustada con los datos de entrenamiento.

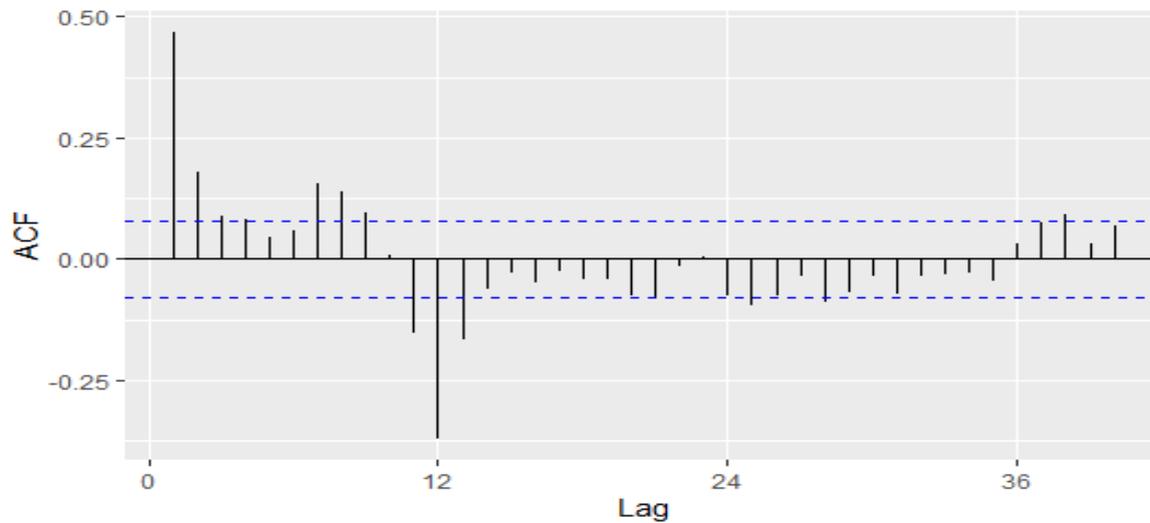
Se realizaron diversas pruebas con diferentes Inputs para esta red, probando un total de 6 modelos diferentes para medir sus resultados. Siguiendo a Pedroza Robles (s/a) se ajustaron 3 modelos escalando los Inputs de la red ($NNAR(1,1,2)_{12}$, $NNAR(11,3,8)_{12}$, $NNAR(11,3,7)_{12}$), así mismo motivados por los resultados obtenidos por Gheyas y Leslie Smith (2009) se ajustaron 3 modelos los cuales antes se volvieron estacionarios²⁴, ($NNAR(11,2,7)_{12}$, $NNAR(11,2,7)_{12}$, $NNAR(11,3,7)_{12}$) a uno de ellos de igual modo se le aplicó una transformación logarítmica ($NNAR(11,2,7)_{12}$), así mismo se probó la eficiencia de la red con este modelo haciéndolo estacionario pero sin aplicar una transformación logarítmica.

El número de retardos para el modelo $NNAR(1,1,2)_{12}$ se escogió arbitrariamente, sin embargo, para el resto de los modelos que se ajustaron se empleó la función de autocorrelación, así como también la función de autocorrelación parcial. Para el modelo

²⁴ Se aplicaron primeras diferencias para eliminar el efecto de la tendencia, así como diferencias estacionales para suavizar el comportamiento estacional del conjunto de datos.

NNAR(11,2,7)₁₂, al cual se le aplicó una transformación logarítmica, la función de autocorrelación y la función de autocorrelación parcial de este pueden verse en los Gráficos 4.11 y 4.12 respectivamente. Para definir el número de parámetros de la red de los demás modelos se muestra a continuación el ACF²⁵ y el PACF²⁶ del proceso diferenciado eliminando los efectos de la tendencia (primeras diferencias) y suavizando el efecto estacional (diferencias estacionales).

Gráfico 4.72. Funciones de Autocorrelación del proceso diferenciado completamente aplicado al conjunto de entrenamiento sin transformación logarítmica.

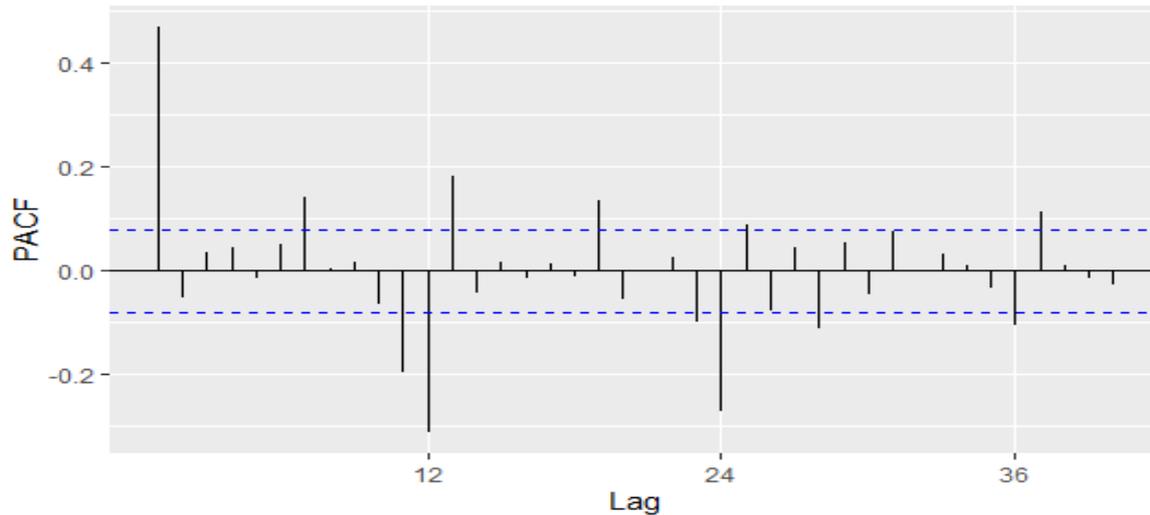


Fuente: Elaboración propia.

²⁵ Función de Autocorrelación

²⁶ Función de Autocorrelación Parcial

Gráfico 4.73. Funciones de Autocorrelación Parcial del proceso diferenciado completamente aplicado al conjunto de entrenamiento sin transformación logarítmica.



Fuente: Elaboración propia.

Como puede apreciarse en los Gráficos 4.72 y 4.73 pueden considerarse los primeros 11 rezagos para armar la red neuronal, así mismo pueden incluirse los efectos estacionales de hasta 3 periodos anteriores. A continuación, en las tablas 4.26 y 4.27 se muestran a manera de resumen el tipo de tratamiento que se le dio a cada modelo de red correspondiente.

Tabla 4.26. Resumen de los modelos de Red Neuronal Ajustados al conjunto de datos

Modelo	Transformación	Escalamiento de datos	Estacionario	Rezagos
<i>NNAR (1, 1, 2)₁₂</i>	Ninguna	Si	No	1
<i>NNAR (11, 3, 8)₁₂</i>	Ninguna	Si	No	11
<i>NNAR (11, 3, 7)₁₂</i>	Ninguna	Si	No	11
<i>NNAR (11, 2, 7)₁₂</i>	logarítmica	No	Si	11
<i>NNAR (11, 2, 7)₁₂</i>	Ninguna	No	Si	11
<i>NNAR (11, 3, 7)₁₂</i>	Ninguna	No	Si	11

Fuente: Elaboración propia.

Tabla 4.27. Resumen de los modelos de Red Neuronal Ajustados al conjunto de datos

Modelo	Neuronas²⁷
<i>NNAR (1, 1, 2)₁₂</i>	2
<i>NNAR (11, 3, 8)₁₂</i>	8
<i>NNAR (11, 3, 7)₁₂</i>	7
<i>NNAR (11, 2, 7)₁₂</i>	7
<i>NNAR (11, 2, 7)₁₂</i>	7
<i>NNAR (11, 3, 7)₁₂</i>	7

Fuente: Elaboración propia.

El paso siguiente consistió en realizar los pronósticos para todo el año 2021 y probar la eficiencia respecto al conjunto de prueba, es aquí en donde se explica el proceso de reversión para las redes que fueron estacionarias. Como los datos de entrada para la red fueron diferenciados en primer orden y diferenciados estacionalmente, el pronóstico arrojado por la red fue en términos de diferencias, por lo que para revertir el proceso solamente se suma a este el resultado de las diferencias estacionales, matemáticamente se tiene lo siguiente:

$$(\mathbf{1} - \mathbf{B}^{12})(\mathbf{1} - \mathbf{B})x_t = (x_t - x_{t-1}) - (x_{t-12} - x_{t-13}) \quad (3)$$

Como el lado izquierdo de (3) es la parte pronosticada por la red, por comodidad se expresará ese lado de la igualdad como \hat{y}_t , luego entonces podemos pasar del mismo lado el resultado de las diferencias estacionales, quedando entonces la siguiente expresión:

$$\hat{y}_t + (x_{t-12} - x_{t-13}) = (x_t - x_{t-1}) \quad (4)$$

Con (4) se obtiene el resultado de las primeras diferencias, por ende, para obtener el valor real de interés, que en este caso es el pronóstico del INPC, basta con pasar de igual modo el valor del mes anterior de la variable, que en este caso es x_{t-1} , quedando entonces la ecuación (5) que es el despeje de la variable de interés.

²⁷ Numero de neuronas que intervienen en la capa oculta

$$\hat{x}_t = \hat{y}_t + (x_{t-12} - x_{t-13}) + x_{t-1} \quad (5)$$

Cabe señalar que este mismo proceso se aplicó indistintamente para aquellas redes cuyos inputs fueron estacionarios de primera instancia, si a los datos de entrada de la red en cuestión se les aplicó una transformación logarítmica, lo único adicional que se implementó fue revertir dicha transformación con la función exponencial. Una vez descritos los procesos de reversión para los modelos cuyos Inputs fueron estacionarios, a continuación, se muestran los resultados para el pronóstico de todo el año 2021 (tablas 4.28, 4.29 y 4.30).

Tabla 4.28. Valores Pronosticados del INPC

Fecha	Real	$NNAR(11, 2, 7)_{12}$ Estacionario-Log	$NNAR(1, 1, 2)_{12}$ Escalado
Ene 2021	110.210	109.703	109.441
Feb 2021	110.907	109.808	109.620
Mar 2021	111.824	109.694	109.778
Abr 2021	112.190	109.134	109.855
May 2021	112.419	109.672	109.949
Jun 2021	113.018	110.174	110.067
Jul 2021	113.682	110.708	110.214
Ago 2021	113.899	111.016	110.372
Sep 2021	114.601	111.209	110.527
Oct 2021	115.561	111.704	110.705
Nov 2021	116.884	112.076	110.869
Dic 2021	117.308	112.812	111.040

Fuente: Elaboración propia.

Tabla 4.29. Valores Pronosticados del INPC

Fecha	Real	<i>NNAR (11, 3, 8)₁₂</i>	<i>NNAR (11, 3, 7)₁₂</i>
		Escalado	Escalado
Ene 2021	110.210	109.610	109.651
Feb 2021	110.907	109.774	109.843
Mar 2021	111.824	109.672	109.794
Abr 2021	112.190	109.681	109.875
May 2021	112.4190	110.067	110.293
Jun 2021	113.018	110.444	110.683
Jul 2021	113.682	110.758	111.012
Ago 2021	113.899	111.111	111.399
Sep 2021	114.601	111.573	111.867
Oct 2021	115.561	111.894	112.188
Nov 2021	116.884	112.131	112.472
Dic 2021	117.308	112.392	112.794

Fuente: Elaboración propia.

Tabla 4.30. Valores Pronosticados del INPC

Fecha	Real	<i>NNAR (11, 2, 7)₁₂</i>	<i>NNAR (11, 3, 7)₁₂</i>
		Estacionario	Estacionario
Ene 2021	110.210	109.610	109.631
Feb 2021	110.907	110.483	110.242
Mar 2021	111.824	110.642	110.693
Abr 2021	112.190	109.930	110.098
May 2021	112.419	110.370	110.460
Jun 2021	113.018	110.565	110.628
Jul 2021	113.682	111.056	111.131
Ago 2021	113.899	111.414	111.390
Sep 2021	114.601	111.758	111.639
Oct 2021	115.561	112.449	112.292
Nov 2021	116.884	112.918	112.837
Dic 2021	117.308	113.691	113.801

Fuente: Elaboración propia.

Así mismo en las tablas 4.31 y 4.32 se muestran algunas medidas de error calculadas con los pronósticos de cada red.

Tabla 4.31. Medidas de error de las redes ajustadas

Medida de error	<i>NNAR (11, 2, 7)₁₂</i>	<i>NNAR (1, 1, 2)₁₂</i>	<i>NNAR (11, 3, 8)₁₂</i>
	Estacionario-Log	Escalado	Escalado
MAE	2.899	3.339	2.783
RMSE	3.134	3.728	3.033
MAPE	0.025	0.029	0.024

Fuente: Elaboración propia.

Tabla 4.32. Medidas de error de las redes ajustadas

Medida de error	<i>NNAR (11, 3, 7)₁₂</i>	<i>NNAR (11, 2, 7)₁₂</i>	<i>NNAR (11, 3, 7)₁₂</i>
	Escalado	Estacionario	Estacionario
MAE	2.553	2.301	2.305
RMSE	2.783	2.532	2.532
MAPE	0.022	0.020	0.020

Fuente: Elaboración propia.

Como se aprecia de todos los posibles candidatos, el modelo de Red Neuronal que comete menos errores es el modelo *NNAR(11,2,7)₁₂* sin transformación logarítmica y volviendo estacionario el conjunto de datos de entrada, por lo que dicho modelo fue el utilizado al momento de pronosticar el comportamiento del INPC para todo 2022. Los resultados de pronóstico para este tipo de red se resumen a continuación en la Tabla 4.33, así mismo se muestra la inflación mensual anualizada obtenida con (2) de la sección 4.1.

Tabla 4.33. Pronósticos del INPC e Inflación mensual anualizada para 2022

Fecha	Pronóstico del INPC	Inflación Mensual Anualizada
Ene 2022	118.428	7.460%
Feb 2022	119.220	7.500%
Mar 2022	119.929	7.250%
Abr 2022	119.791	6.770%
May 2022	120.534	7.220%
Jun 2022	121.386	7.400%
Jul 2022	121.959	7.280%
Ago 2022	122.045	7.150%
Sep 2022	122.752	7.110%
Oct 2022	123.968	7.270%
Nov 2022	125.250	7.160%
Dic 2022	125.787	7.230%

Fuente: Elaboración propia utilizando (3) para revertir el proceso y (2) de la sección 4.1 para obtener la inflación mensual anualizada en México.

Como se puede apreciar, de igual modo la red neuronal otorga pronósticos altos, confirmando de igual modo la persistencia de la inflación en México ocasionada por los factores anteriormente descritos en los capítulos previos de esta tesis. Visualmente se muestra en el Gráfico 4.74 la comparación de dichos valores a través del tiempo, confirmando que los valores van más allá del objetivo inflacionario en México que es de a lo más 4%.

Gráfico 4.74. Inflación mensual anualizada pronosticada con el modelo $NNAR(11, 2, 7)_{12}$.



Fuente: Elaboración propia con los pronósticos de la red $NNAR(11,2,7)_{12}$ y Excel 365 para generar el gráfico.

Discusión y Conclusiones

Una vez expuestos los resultados de este proyecto, en este capítulo se presenta el siguiente análisis de acuerdo con el desempeño de cada uno de los modelos construidos y lograr los pronósticos del Índice Nacional de Precios al Consumidor que fue el objetivo de esta tesis.

En orden del desarrollo de los modelos propuestos, para el modelo SARIMA que se construyó en el apartado 4.1 destaca que los pronósticos del INPC son adecuados pese al número de parámetros resultantes en la ecuación elegida. Si bien, se buscó que estos fueran mínimos, adaptándose al crecimiento que tiene esta variable mes a mes durante un año en particular, sin embargo, por temas económicos a nivel global el modelo por sí solo no logra adaptarse tan bien a dicha volatilidad. Es por ello que surgió la necesidad de implementar una herramienta adicional que ayudara a los pronósticos, esta fue la construcción de intervalos de confianza construidos con base en un modelo GARCH(1,1) esto con la finalidad de tener un amplio panorama del posible comportamiento de la variable de interés (INPC), así como también poder adaptarle la volatilidad implícita de la serie al modelo y pueda captar dicho componente. Los resultados de tal procedimiento mostraron que al implementar esta herramienta de pronóstico los valores logran tener un mejor desempeño en cuanto a la precisión o al menos proporciona información adicional sobre lo que podía suceder durante el periodo elegido para pronóstico (2022) que como se pudo ver en los pronósticos, se esperaba una inflación superior al objetivo establecido por el Banco de México, confirmando así un periodo inflacionario más dentro de la historia de nuestro país.

Por otro lado, el modelo SARIMAX con variables exógenas resultó adecuado para pronosticar al INPC, pues se demostró que las variables elegidas además de tener el respaldo teórico se encuentran estadísticamente relacionadas con la variable dependiente. Además, se encontraron modelos en donde todas las variables fueron significativas. Sin embargo, se descartaron porque se encontraron otros con menos variables que cumplían con tener las medidas de error más bajas y sobre todo eran más eficaces al momento de realizar los pronósticos (al compararlos con los valores reales

publicados para 2022). A pesar de que es recomendable tener muchos datos en las series de tiempo, el número contemplado en este modelo fue adecuado y suficiente.

Las variables incluidas en los modelos finales (salario mínimo, oferta y demanda agregada y expectativas inflacionarias) son adecuadas para continuar con los pronósticos de la inflación en los próximos años, ya que son variables generales que se seguirán midiendo y en el caso de las expectativas arrojadas por la encuesta realizada a los expertos incluirán el contexto internacional, lo que reduce la necesidad de recurrir a otras variables que puedan llegar a influir en el comportamiento de la inflación.

Así mismo para cada modelo de Red Neuronal el aprendizaje resultó ser muy diferente dependiendo de los parámetros que se le asigne a cada red y más importante aún se detectó una posible dependencia sobre el número de neuronas de entrada con la que se alimente ya que al final del día el funcionamiento de una red neuronal se basa en las diferentes combinaciones de los datos de entrada para ajustar el valor de salida.

Centrándose en las primeras redes neuronales (sección 4.3.1, 4.3.2, 4.3.3 y 4.3.4), que se llevaron a cabo en Python, se pudo apreciar que algunos de los modelos hicieron ajustes bastante buenos en las primeras épocas o iteraciones del aprendizaje. Sin embargo, la precisión de cada modelo fue muy baja ya que los datos analizados difícilmente se repetirán en la salida de la red por el hecho de que se están haciendo ajustes con los valores decimales e incluso se correría el riesgo de un sobre entrenamiento en el caso de que la precisión incrementa.

Así mismo se puede ver que en todos los ajustes propuestos en inicio subestiman el valor real publicado para 2022, truncándose incluso en el último en un valor por más menos 110 en cuanto al INPC. En resumidas cuentas, estos modelos por el tipo de red ajustada no logran aprender del todo el comportamiento de la variable de interés. Cabe resaltar que como en toda red neuronal el pronóstico dependerá de los datos con los que se alimente ya que entre más datos puedan considerarse o en su defecto se escoja el tratamiento adecuado para los mismos en la capa de entrada el ajuste será mejor. Es por ello que se optó por otra alternativa de pronóstico utilizando redes neuronales autorregresivas, la cual

como bien se describió en la sección 4.3.5 utiliza rezagos y rezagos múltiples de 12 (por ser datos con estacionalidad) de la variable de interés para predecir su comportamiento en el periodo de pronóstico elegido (2022). Sobre esto, se puede comentar que analizando diversos trabajos Gheyas y Smith (2009) se encontraron formas diferentes para entrenar este tipo de red: introduciendo datos escalados o haciendo estacionario el conjunto de datos disponible (Pedroza Robles, s/a).

En conclusión, para este trabajo el modelo de red neuronal que arroja mejores pronósticos es la red neuronal autorregresiva ($NNAR(11,2,7)_{12}$) sin transformación logarítmica y volviendo estacionario el conjunto de datos de entrada, confirmando los resultados propuestos por Gheyas y Smith (2009) respecto al mejoramiento de los pronósticos.

Se confirma de igual modo que influye bastante el cómo se alimente la red a través de las neuronas de entrada ya que a diferencia de los primeros modelos de redes ajustados, para este último, el pronóstico del INPC, así como el de la inflación anualizada se pudo predecir eficientemente minimizando las medidas de error pertinentes. Cabe destacar que las bondades de usar las nuevas metodologías de pronóstico (red neuronal autorregresiva) ayudan bastante a la limitante que proponen los modelos de serie de tiempo respecto a la estacionariedad y/o número de parámetros a tomar en cuenta, ya que estas metodologías otorgan una mejor aproximación sin tantas restricciones.

Limitaciones de la tesis

Dentro de algunas limitaciones importantes para este trabajo destaca la mejora de algunos métodos de pronóstico:

- 1) Para el modelo SARIMA-GARCH puede ser que el análisis solo se centró en las variaciones positivas de la variable respecto a la volatilidad de sus residuales pueda ser un problema ya que usualmente los modelos GARCH solo contemplan los cambios positivos de la volatilidad. Se propone entonces, considerar otros métodos como el E-GARCH (que toma en cuenta cambios positivos y negativos) para modelar este componente puedan mejorar los pronósticos establecidos en este trabajo.

- 2) Por otro lado, la disposición de los datos en todas las variables que deseen incluirse en el modelo es la mayor limitante presente dentro del modelo SARIMAX, ya que como se vio en esta investigación se deben adaptar los datos, la temporalidad tiene que ser la misma (quincenal, mensual, diaria, etc.), y deben ser el mismo número de observaciones. En este sentido, la presencia de una variable de la que no se tienen muchos registros no es adecuada para este tipo de modelo (como sucedió con la variable COVID, que no se pudo incluir en este trabajo, si bien se sabe que la enfermedad tuvo un papel importante dentro de la economía mexicana), por lo que las variables disponibles siempre podrán verse reducidas al utilizarse este modelo.
- 3) Continuando con la homogeneidad entre las variables es importante tener en cuenta la estacionalidad que puedan presentar las mismas, en esta tesis para variable dependiente se tuvo siempre presente el problema de estacionalidad, por lo que todas las variables exógenas fueron tratadas de la misma forma. Sin embargo, si la variable dependiente elegida no tiene ese problema y alguna de las independientes si, no se aplicarán las diferencias correspondientes y no se podrá conseguir que todos los procesos sean estacionarios, reduciendo así las variables que pudieran ser utilizadas para construir el modelo.
- 4) En este caso, el modelo SARIMAX implementado no puede ser comparable con los otros dos modelos desarrollados debido a que no se estudia el mismo periodo de tiempo, por lo que en esta investigación queda más como un ejercicio de cómo plantear esta variante del modelo ARIMA.
- 5) Una de las principales limitantes que se tuvieron para este trabajo fue el hecho de que actualmente existen situaciones atípicas a nivel global derivadas de lo que fue la pandemia por el virus SARS-COV2, así como los conflictos bélicos entre Rusia y Ucrania, los cuales dejaron un impacto económico que ha llenado de incertidumbre los mercados y las principales economías globales. La economía sufrió una parálisis total al inicio de la pandemia, ocasionando que las principales economías empezaran a caer y los precios de ciertos bienes incluso llegaran a precios nunca imaginados, ocasionando que al pasar el tiempo se empezó a tener cierta incertidumbre de la duración de esta misma parálisis económica, así como el hecho de que las personas han empezado a volver a sus vidas normales.

6) Otro elemento a considerar es el hecho de reactivar la economía de un momento a otro después de meses de actividad baja, nos ha llevado a un escenario donde las tasas inflacionarias de diferentes países se disparara y llegara a porcentajes no reportados durante bastante tiempo como lo fue en el país y se habla de otros países ya que como se sabe, México es una economía sensible a los cambios sobre la economía global llevándonos a un escenario de incertidumbre total donde expertos en el tema financiero han pronosticado una recesión importante para el país seguida de una recuperación e incluso se habla de una estanflación. Así, pese que los modelos predecían una tasa de inflación alta para los meses siguientes del año elegido para pronostico (2022), esos pronósticos puedan verse bastante sesgados por la incertidumbre actual que se vive por el bajo poder adquisitivo que ha dejado la pandemia a la población mexicana siendo esta una limitante para el trabajo ya que no se considera dentro de él los efectos directos de esta pandemia o las situaciones que dicho conflicto ha dejado de forma directa o indirecta.

Futuras líneas de investigación

Como pudo verse, estos modelos resultaron prácticos al momento de llevar a cabo los pronósticos. Sin embargo, estos pueden retomarse a futuro para ver si con la información adicional que se tiene sobre esta crisis global se pueda tener una mejor aproximación a los valores reales. Por otro lado, y debido a la característica presentada en esta serie, pueden ajustarse otros modelos que entiendan de mejor manera el comportamiento de los datos, puede ser el caso de la inclusión de un modelo ARFIMA debido al problema de memoria larga que presenta esta serie, así mismo se podría incluir a los modelos de regresión dinámica debido a que permite incluir más variables que puedan explicar el comportamiento de la inflación. Como se mostró en el apartado 4.2 los modelos finales arrojaron pronósticos acertados y muy cerca de los valores reales de los que se tiene registro este año. Sin embargo, el modelo podría complementarse construyendo intervalos de confianza que permitan tener un panorama más amplio de los valores inferiores o superiores que pueda alcanzar la inflación.

Incluir la volatilidad de la serie en un modelo SARIMA-GARCH también puede ser una buena alternativa para mejorar los pronósticos, ya que a pesar de haber eliminado las

partes más volátiles de la serie la prueba ARCH para los datos acortados rechaza la hipótesis de que se haya eliminado esta característica.

Anexos

Tabla A.1. Prueba de estacionariedad para el INPC en México

KPPS	P-VALUE	DICKEY-FULLER	P-VALUE
8.712	0.010	-1.481	0.798

Elaboración propia con las funciones `kpps.test()` y `adf.test()` en RStudio.

Tabla A.2. Prueba de estacionariedad para el INPC en México diferenciado completamente.

KPPS	P-VALUE	DICKEY-FULLER	P-VALUE
0.082	0.100	-6.956	0.010

Elaboración propia.

Tabla A.3. Resumen estadístico para los coeficientes del modelo GARCH (1,1) ajustado a los residuales de la ecuación candidata.

Coeficiente	Estimate	T	p-value
Omega	0.000	3.563	0.000
alpha1	0.375	6.871	0.000
beta1	0.624	31.986	0.000

Fuente: Elaboración propia.

Tabla A.4. Prueba de estacionariedad del INPC

KPPS	P-VALUE	DICKEY-FULLER	P-VALUE
4.50	0.010	2.284	0.990

Elaboración propia.

Tabla A.5. Prueba de estacionariedad del INPC conjunto de entrenamiento con transformación y primeras diferencias.

KPPS	P-VALUE	DICKEY-FULLER	P-VALUE
0.122	0.100	-8.459	0.010

Elaboración propia.

Tabla A.6. Prueba de estacionariedad de las variables exógenas

Variable	KPPS	P-VALUE	DICKEY-FULLER	P-VALUE
Tipo de cambio	0.020	0.100	-6.000	0.010
Salario mínimo	0.102	0.100	-6.730	0.010
Expectativas inflacionarias	0.018	0.100	-10.871	0.010
Oferta y demanda agregada	0.031	0.100	-6.311	0.010

Fuente: Elaboración propia

Bibliografía

- Banco de México. (2011). *Informe Anual 2010*.
- Banco de México. (2012). *Informe Anual 2011*.
- Banco de México. (2013). *Informe Anual 2012*.
- Banco de México. (2014). *Compilación de informe trimestral*.
- Banco de México. (2015). *Compilación de informe trimestral*.
- Banco de México. (2016). *Compilación de informe trimestral*.
- Banco de México. (2017). *Compilación de informe trimestral*.
- Banco de México. (2018). *Compilación de informe trimestral*.
- Banco de México. (2019). *Compilación de informe trimestral*.
- Banco de México. (2020). *Compilación de informe trimestral*.
- Banco de México. (2021). *Compilación de informe trimestral*.
- Banco de México. (2022). *Compilación de informe trimestral*.
- BANXICO. (junio de 2016). *Salario mínimo e inflación*.
- BANXICO. (Agosto de 2018). *Principales Elementos del Cambio base del INPC*. México. Obtenido de banxico.org.mx: <https://www.banxico.org.mx/publicaciones-y-prensa/informes-trimestrales/recuadros/%7B1433DE85-D1A1-672C-CAF2-17E95DBA5BC0%7D.pdf>
- BANXICO. (2019). *Evolución de las Expectativas de Inflación de Largo Plazo*. Estracto del informe trimestral Octubre-Diciembre.
- Banxico. (2020). *ABC DE LA POLITICA MONETARIA*. Mexico.
- BANXICO. (2020). *Impacto de la Pandemia de COVID-19 en la Dinámica de Precios al Consumidor en EUA y México*.
- BANXICO. (2021). *Factores Asociados a la Evolución Reciente de la Inflación en Estados Unidos*.
- BANXICO. (2021). *Informe trimestral Abril- Junio 2021, Resumen visual*.
- Banxico. (s.f.). *Banxico*. Obtenido de Informe trimestral: Enero-Marzo 2022: <https://www.banxico.org.mx/publicaciones-y-prensa/informes-trimestrales/informes-trimestrales-precios.html>
- Banxico. (s.f.). *Banxico-Educa*. Obtenido de <http://educa.banxico.org.mx/economia/inflacion-economia.html>
- Blanchard, O., Amighini, A., & Giavazzi, F. (2012). *Macroeconomía*. Madrid: Pearson.

- Calvo, J. (20 de Julio de 2020). *Crear Red Neuronal desde las matematicas*. Obtenido de European Valley: <https://www.europeanvalley.es/noticias/crear-red-neuronal-desde-las-matematicas/>
- Cantú, J. (2018). Variación en el costo del peso mexicano ante el dólar estadounidense. *Revista Espacios*, 20.
- Carmona, F., Montaña, G, Carrion, J, & Aguilar, A. (1970). *El milagro mexicano*. Ciudad de México: Editorial nuestro tiempo S.A.
- Carsten, T., & Gándara, G. (1990). El plan Brady y la negociación de la deuda externa. *Comercio Exterior*, 303-308.
- Collado, M. (2011). Autoritarismo en tiempos de crisis. Miguel de la Madrid 1982-1988. *Historia y grafía*, 150-177.
- CONASAMI. (Julio de 2011). *conasami.gob.mx*. Obtenido de http://www.conasami.gob.mx/pdf/participacion_ciudadana/PARTICIPACION_CIUADANA_JULIO_2011.pdf
- CONSAR. (2021). *Impacto de la Pandemia de Covid-19 en el mercado laboral Mexicano y en el SAR*. Ciudad de México.
- Contreras, J. C., Santana Contreras, J. C., Camaro, A. A., Casas Henao, A., & Jimenez Mendez, E. (28 de Diciembre de 2006). *Scielo*. Obtenido de Scielo: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-50512006000200011
- Delgadillo, O., Ramírez, P., leos, J., Salas, J. M., & Valdez, R. (2016). Pronósticos y series de tiempo de rendimientos de granos básicos en México. *Acta Universitaria*, 23-32.
- Economía UNAM. (abril de 2010). Estancamiento económico, desigualdad y pobreza: 1982-2009. En C. Tello. México.
- El Economista. (s.f.). *Eleconomista*. Obtenido de Inflación: <https://www.eleconomista.es/diccionario-de-economia/inflacion>
- Engel, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance in the United Kingdom. *Econometrica*, 987-1008.
- Engel, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance in United Kingdom. *Econometrica*, 987-1008.
- FEDESARROLLO. (1974). Impacto de la Crisis petrolera mundial. *Repository Fedesarrollo*.
- Garcia, A. T. (2002). *Un Analisis de las Tasas de Interés en México a través de la Metodología de Reglas Monetarias*. México: Banco de México.
- Gómez, C. y. (agosto de 2017). *Eumed*. Obtenido de <https://www.eumed.net/coursecon/ecolat/mx/2017/devaluaciones-peso-mexico.html>

- Gujarati, D., & Porter, D. (2010). *Econometría*. Mc Graw Hill.
- Heath, J. (2012). En J. Heath, *Lo que indican los indicadores: Cómo utilizar la información estadística para entender la realidad económica de México* (pág. 14). México.
- Heath, J. (2012). *Lo que indican los indicadores*. México: INEGI.
- Heath, J. (2012). *Lo que indican los indicadores*. México: INEGI.
- Hernández, S. (2015). *CURSO REGIONAL SOBRE HOJA DE BALANCE DE ALIMENTOS, SERIES DE TIEMPO Y ANÁLISIS DE POLÍTICA*. Obtenido de cepal.org: https://www.cepal.org/sites/default/files/courses/files/01_2_arima.pdf
- Hyndman, R. (s.f.). *Scale-dependent errors*. Obtenido de Forecasting: Principles and Practice: <https://otexts.com/fpp3/accuracy.html>
- Hyndman, R., & Athanasopoulos, G. (2022). Obtenido de Forecasting: Principles and Practice: <https://otexts.com/fpp3/>
- Hyndman, R., & Athanasopoulos, G. (2022). Obtenido de Forecasting: Principles and Practice: <https://otexts.com/fpp3/prediction-intervals.html>
- IBM. (17 de 08 de 2021). *ibm.com*. Obtenido de <https://www.ibm.com/docs/es/spss-modeler/saas?topic=networks-neural-model>
- INEGI. (s.f.). *INEGI*. Obtenido de ¿A qué se refiere la inflación subyacente y no subyacente?: <https://www.inegi.org.mx/programas/inpc/2018/PreguntasF/>
- Inflación. (2012). En J. Heath, *Lo que indican los indicadores: Cómo utilizar la información estadística para entender la realidad económica de México*. (págs. 293-295). México.
- Informe trimestral. (2022). *Banxico*. Obtenido de <https://www.banxico.org.mx/publicaciones-y-prensa/informes-trimestrales/informes-trimestrales-precios.html>
- López, C. (11 de 09 de 2019). *La relación entre inflación y tipo de cambio en México*. Obtenido de UFM Market Trends: <https://trends.ufm.edu/articulo/inflacion-tipo-cambio-mexico/>
- Massó, V., & Pons, E. (20 de 03 de 2018). *Caminata aleatoria o "Random Walk"*. Obtenido de https://rstudio-pubs-static.s3.amazonaws.com/372908_51086b0fd9af49e3966d372046e7969a.html
- Medina, L. (2006). Los años de Salinas: Crisis electoral y reformas. *CIDE*.
- México, B. d. (s.f.). *Banxico*. Obtenido de Informe trimestral enero-marzo 2022: <https://www.banxico.org.mx/publicaciones-y-prensa/informes-trimestrales/informes-trimestrales-precios.html>

- Monex. (28 de 12 de 2020). *¿Cómo afecta la volatilidad del tipo de cambio a la economía?* Obtenido de Monex: <https://blog.monex.com.mx/como-afecta-la-volatilidad-del-tipo-de-cambio-a-la-economia>
- Morales, I. (1986). *Las negociaciones del gas entre 1977 y 1979*.
- Olivera, O. G.-O. (16 de Septiembre de 2019). *Xeridia*. Obtenido de Xeridia: <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>
- Pengfei, L. (11 de Abril de 2005). *Box-Cox Transformation: An Overview*.
- Peña, R. (1979). La política Económica Mexicana 1970-1976. *Nexos*.
- Rambiola, N. (23 de Diciembre de 2019). *elinpc.com.mx*. Obtenido de <https://elinpc.com.mx/salario-minimo-en-mexico/>
- Romero, J. (2019). Obtenido de core.ac.uk: <https://core.ac.uk/download/pdf/250160185.pdf>
- Sacristán, E. (2006). Las privatizaciones en México. *Economía UNAM*, 54-64.
- Tello, C. (2011). *Notas sobre el Desarrollo Estabilizador*.
- Torres García, A. (2003). *Reglas de la política monetaria como ancla nominal: evidencia de la economía mexicana*. Ciudad de México: Centro de Estudios Monetarios Latinoamericanos.
- Tsay, R. (s.f.). Identifying AR Models in Practice. En R. Tsay, *Analysis of Financial Time Series* (págs. 40-41). Wiley Series in Probability and Statistics.
- Tsay, R. (s.f.). Magnitude Measure. En R. Tsay, *Analysis of Financial Time Series* (pág. 194). Wiley Series in Probability and Statistics.
- UNAM. (2002). Como llegamos a la crisis de 1994.
- Vidal Gonzalez, M. (2015). *El uso del Perceptrón Multicapa para la clasificación de patrones en conductas adictivas*. Palma.
- Villanueva Epinoza, M. d. (2002). Elementos de una red neuronal artificial. En M. d. Villanueva Epinoza, *Las redes neuronales artificiales y su importancia como herramienta en la toma de decisiones*. Lima.
- Whitehead, L. (1980). La política económica del sexenio de Echeverría, ¿qué salió mal y por qué? *Foro Internacional*, 484-515.