

Filtro de Kalman para el pronóstico del desempleo en México (2007M1-2021M12)

Roberto Guillermo De la Cruz Yañez

Resumen

Se realiza la comparación de la eficiencia del filtro de Kalman en la realización de pronósticos sobre la tasa de desempleo mensual para México respecto a metodologías convencionales (ARIMA y VAR) tomando como periodo observado 2007m01 a 2019m03 y pronosticando a partir de 2019m04 hasta 2021m12.

La conclusión principal es que el filtro de Kalman obtiene un coeficiente de Theil menor (0.01) respecto a las otras metodologías (0.13 y 0.03, respectivamente).

Además, se comprueba que el filtro de Kalman realiza un pronóstico más eficiente debido a que tiene la mejor simulación histórica para la tasa de desempleo y posee una dinámica similar a la serie observada.

Palabras clave: Pronósticos, Kalman, Modelos estado-espacio, desempleo.

Código JEL (*Journal of Econmical Literature*): C01, E24, E27, C53

Introducción

La importancia de los pronósticos en la economía mundial y su aplicación en las políticas económicas han vivido un proceso de rápido crecimiento y de implementación de nuevas y más sofisticadas herramientas con el fin de predecir de mejor forma el comportamiento de las principales variables económicas de un país o región.

La econometría se ha convertido en una herramienta esencial para construir predicciones, aunque se ha cuestionado la utilidad de los modelos econométricos especialmente en el contexto de la crisis de 2009 (Ferrari, 2008). Una vía para mejorar la implementación de estas técnicas es el uso de modelos alternativos que, conforme a una evaluación de precisión, busquen seleccionar la técnica que logre optimizar el proceso de predicción. Esta estrategia podría considerarse una buena manera para mejorar la eficiencia de los pronósticos, lo cual es un objetivo importante de los economistas contemporáneos, principalmente dado la incertidumbre en las principales variables macroeconómicas y su inferencia en la toma de decisiones.

La creación de diferentes formas y la implementación de nuevos modelos en la búsqueda de pronósticos más eficientes trajo consigo un problema práctico: dado diferentes metodologías se proporcionan diferentes pronósticos para una misma variable.

Este estudio analiza la eficiencia de un modelo multivariable dinámico construido a partir de la metodología de filtro de Kalman para predecir el comportamiento de la tasa de desempleo en México a partir de mayo de 2019 hasta diciembre del 2021.

La contribución de este estudio radica en mostrar la utilidad del filtro de Kalman en la realización de un pronóstico, a fin de medir su capacidad para predecir el comportamiento de una variable clave en el devenir de la economía. Dicho modelo posee un algoritmo que permite realizar inferencia estadística en un sistema dinámico con características de modelo

bayesiano similares al modelo oculto de Márkov¹ (HMM por sus siglas en ingles) y el cual se resuelve por un modelo de estado-espacio que emplea variables observadas continuas con propiedades de distribución probabilística gaussiana.²

El filtro de Kalman generalmente es aplicado para determinar la tasa de desempleo natural (NAIRU, por sus siglas en ingles). Sin embargo, existen otros enfoques que plantean el uso del filtro de Kalman como una potente herramienta en la elaboración de pronósticos económicos (Bernanke, 1997). Esta metodología ofrece una forma mucho más flexible, permitiendo trabajar con varios modelos, con la ventaja de ir actualizando los datos a medida que arriban los valores observados (Paganini, 2003). También es posible la incorporación de datos externos provenientes de series relacionadas.

La importancia de obtener un correcto pronóstico del desempleo radica en el impacto que tiene esta variable sobre el comportamiento de la economía en general, ya que esta variable puede reflejar de mejor forma el nivel de bienestar que poseen las familias dentro de una región o país (Heath, 2012). Además, que al ser un indicador con frecuencia más alta que otras variables, puede enviar mensajes importantes para los tomadores de decisiones o inversores sobre el desempeño inmediato de otras variables.

El pronóstico de la tasa de desempleo es un tema crucial en la elaboración de planes a futuro, sobre todo por el impacto de esta sobre otras variables macroeconómicas y microeconómicas. Por ende, la exactitud en su elaboración y acierto, debe de ser considerado un aspecto fundamental para la toma de decisiones políticas, instituciones, tomadores de decisión, inversores etc. Para lograr la construcción de este modelo se han tomado como consideración las variables explicativas: remuneraciones medias reales e IGAE, ambas con periodicidad mensual y en diferencias.

¹ En dicho modelo se asume que el sistema a modelar es un proceso de Márkov de parámetros desconocidos (Simionescu, 2015).

² Esto se cumple cuando la función de densidad de una variable aleatoria corresponde con la distribución normal de media $\mu = b$ y varianza $\sigma^2 = c^2$.

A pesar de la existencia de muchos estudios sobre la evaluación de diferentes metodologías en la elaboración de pronósticos por métodos econométricos, solo algunos de ellos se centran en temas de desempleo y la comparación de metodologías aplicadas a los pronósticos. Camba-Mendez (2012) construyó pronósticos condicionales usando Modelos VAR y técnicas de filtro de Kalman, encontrando que el filtro de Kalman tiene interesantes conclusiones en la elaboración de pronósticos macroeconómicos. Kishor y Koenig (2012) encuentra la ineficiencia de las técnicas VAR en los pronósticos de datos sujetos a revisión, concluyendo que en los modelos VAR para pronosticar la tasa de desempleo genera pronósticos propensos a revisiones posteriores, debido al sesgo en la reconstrucción de la serie que presentan. Sermpinis (2013) realizan predicciones para la tasa de desempleo en los Estados Unidos comprobando que las utilidades de algoritmos adaptativos son más eficientes que los modelos VAR.

Milas y Rothman (2008) utilizaron modelos de corrección de errores de vector de transición suave (STVECM por sus siglas en inglés) para predecir la tasa de desempleo en numerosos países; para la mayoría de países las previsiones de STVECM superaron las predicciones por los métodos convencionales, Proietti (2003) comparó la precisión de varias predicciones basadas en componentes lineales no observados, aplicadas a la tasa de desempleo mensual en los Estados Unidos, concluyendo que los choques no son persistentes durante el ciclo económico.

La organización de esta investigación es la siguiente: después de una breve revisión de la literatura que presenta los métodos cuantitativos utilizados para predecir la tasa de desempleo, se explicará las tres metodologías a comparar. Las predicciones se realizan para el desempleo en México de 2007 a 2021 utilizando modelos ARIMA, VAR y filtro de Kalman, así como comparando la eficiencia de cada una de ellas. La evaluación del pronóstico se basa en el coeficiente de Theil (0.13, 0.03 y 0.01 respectivamente), concluyendo que el filtro de Kalman recrea de mejor forma la simulación histórica para la tasa de desempleo, lo cual indica que esta última metodología es la más adecuada para este ejercicio.

1. Aspectos Teóricos

La importancia de la tasa de desempleo radica en ser un indicador clave para evaluar el desempeño económico del país, de aquí el interés de analizar el papel que juega en la economía mexicana. Sin embargo, resulta complejo la elección de un modelo econométrico parsimonioso y que logre capturar de forma eficiente, por medio de un equilibrio simple entre teoría y relaciones empíricas, el proceso generador de información detrás de la variable seleccionada.

Debido a la importancia económica y social de esta variable, es necesario elegir una metodología que pondere las propiedades de la serie y logre maximizar la elección de las variables que pueden impactar en el desempleo y ocuparlas para pronósticos eficientes.

1.1 Modelos ARIMA

Los modelos Autorregresivos Integrados de Media Móvil (ARIMA por sus siglas en inglés) representan una buena opción para el pronóstico de indicadores de cuentas nacionales con ausencia de información³ (FMI, 2017).

Dado que la palabra ARIMA está integrada por una serie de conceptos importantes para entender las propiedades de esta metodología,⁴ se requiere diseccionar cada una de sus partes para un correcto análisis. Esta metodología consiste en la combinación de un término autorregresivo (AR), un término de promedio móvil (MA) y la diferenciación de la serie (I). La forma de representarlos es por medio de ARIMA (p,d,q); donde (p) se refiere al orden del modelo autorregresivo; (d), al término de diferenciación, y (q), al término de media móvil (Amarís, 2017).

³ Es decir, que la variable a pronosticar carezca de variables explicativas que ayuden a la elaboración de un pronóstico.

⁴ Para la explicación de cada uno de los procesos se ocupa como referencia el texto de Mahia (2003).

Modelos autorregresivos (AR)

Si el comportamiento de una variable en un periodo t es explicada por las observaciones de ella misma correspondientes a periodos anteriores t_{-1} añadiéndose un término de error (Mahía, 2003), se le conoce como un proceso autorregresivo (AR). De tal forma, que una variable Y_t se puede expresar como una combinación lineal de sus valores pasados más una innovación o error.

Para lograr identificar el número de periodos pasados que pueden explicar el comportamiento de la serie el orden del AR será muy importante, ya que este expresa el número de observaciones rezagadas que intervienen en la ecuación. Por tanto, siguiendo la explicación de Mahia (2013), tenemos que:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + a_t \quad (1)$$

donde cada Φ es la información explicativa agregada en cada uno de los periodos de la variable Y , teniendo un proceso AR(p), y en donde el error se le denomina ruido blanco.⁵ De igual forma, puede ser rescrito como:

$$\Phi_p(L)Y_t = \Phi_0 + a_t \quad (2)$$

en donde $\Phi_p(L)Y_t$ se le conoce como el operador polinomial de retardos:

$$\Phi_p(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p \quad (3)$$

y L es lo que se conoce como operador retardo tal que, aplicado al valor de una variable en t , dé como resultado el valor de esa misma variable en t_{-1} :

$$L^p Y_t = Y_{t-p} \quad (4)$$

⁵ Se dice que el error es ruido blanco cuando cumple con: media cero, varianza constante y covarianza nula entre los errores.

Modelo de Media Móvil (MA)

En estos modelos el valor de una variable en un período t , se encontrará en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes (Mahía, 2003), en donde la ponderación de estos valores vendrá determinada por lo cercano o lejano que esté del periodo t . La forma para identificar el orden de estos modelos será igual que en los modelos AR, por lo que se denotan como MA(q) y están definidos por:

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \dots + \theta_q a_{t-q} + a_t \quad (5)$$

en donde μ hace referencia a un término constante, a_t a un término de error tipo ruido blanco y θ_q a la sucesión de errores correspondientes a periodos anteriores. Además, abreviando:

$$Y_t = \theta_q(L)a_t + \mu \quad (6)$$

De tal, forma que, todo proceso estocástico estacionario que puede ser expresado como un modelo del tipo AR(p) podrá ser expresado como un MA(q), llegando a un modelo tipo:

$$Y_t = a_t + \Phi a_{t-1} + \Phi^2 a_{t-2} + \Phi^3 a_{t-3} + \dots + \Phi^j a_{t-j} \quad (7)$$

Estacionariedad e invertibilidad (I)

Una condición fundamental para la representación de un proceso estocástico como un modelo AR o MA, se encuentra restringida al cumplimiento, la menos de forma débil, de la condición de estacionariedad de una serie (Mahía, 2003). Esta será representada por la letra I y va a ser seguida por el orden que esta tiene $I(d)$

Se dice que un proceso estocástico es estacionario en sentido estricto si las funciones de distribución conjunta son invariantes respecto al desplazamiento que esta tiene en el tiempo. Es decir:

$$F(Y_t, Y_{t+1}, \dots, Y_{t+k}) = F(Y_{t+m}, Y_{t+1+m}, \dots, Y_{t+k+m}) \quad (8)$$

Sin embargo, este concepto es sumamente restrictivo, por lo que un acercamiento más real al desempeño de una serie es el buscar estacionariedad débil, y debe de cumplir con:

- a) Las esperanzas matemática no dependen del tiempo, son constantes $E[Y_t] = E[Y_{t+m}] \forall m$
- b) Las varianzas no dependen del tiempo y son finitas $Var[Y_t] = Var[Y_{t+m}] \neq \infty \forall m$
- c) Las covarianzas entre dos variables aleatorias del proceso correspondiente a periodos distintos de tiempo solamente dependen del lapso de tiempo transcurrido entre ellas $Cov(Y_t, Y_s) = Cov(Y_{t+m}, Y_{s+m}) \forall m$

Por tanto, un proceso será integrado de orden d , si la serie original no es estacionaria pero su diferencia de orden d , $\Delta^d Y_t$, sigue un proceso ARMA (p-d,q) estacionario e invertible (González, 2009). En donde, el orden de integración del proceso será las veces necesarias que este se debe de diferenciar para conseguir estacionariedad en la media, es decir, el número de raíces unitarias que tiene el proceso.

Considerando estos tres conceptos definidos se puede afirmar que un modelo ARIMA (p,d,q) estará constituido por el orden del polinomio autorregresivo estacionario p , el orden de integración de la serie d y el orden del polinomio de medias móviles invertible q . Quedando expresado como:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_{p+d} Y_{t-p-d} + \delta + U_t + \theta_1 U_{t-1} + \dots + \theta_q U_{t-q} \quad (9)$$

Es decir:

$$\Phi_p(L)\Delta^d Y_t = \delta + \theta_q(L)a_t \quad (10)$$

donde tanto el polinomio autorregresivo estacionario $\Phi_p(L)$ y polinomio invertible de medias móviles $\theta_q(L)$ no tienen raíces comunes (González, 2009).

1.2 Modelos de Vectores Autorregresivos (VAR)

La metodología de vectores autorregresivos (VAR por sus siglas en inglés) representó una gran revolución en la econometría, conjuntando en un solo modelo herramientas estadísticas y teóricas. Loría (2007) realiza una interesante reflexión sobre la naturaleza de estos modelos, asumiendo que su creación obedece a un sentido epistemológico inverso, dado que le otorgó el peso preponderante a la naturaleza dinámica (memoria) de los datos, pero conforme ha ido evolucionado le ha dado un mayor peso a la teoría económica y a las pruebas de correcta especificación, con el propósito de que se obtengan estimaciones más robustas.

De tal forma, esta metodología proporciona muy interesantes propiedades que pueden capturar efectos sobre las series de tiempo y explotar sus propiedades dinámicas. Una de estas propiedades es la posibilidad de lograr que todas las variables sean de carácter endógeno,⁶ lo cual recupera la naturaleza dinámica de las series de tiempo (Loría, 2007). Es decir, en este modelo cada una de las variables pues cada una se expresa como una función lineal de sus propios valores rezagados y de los valores rezagos de las variables que integran la especificación del modelo, permitiendo analizar la dinámica de la relación de las variables en el corto plazo, lo cual no es detectable con modelos univariados como los ARIMA⁷ (Arias, 2004).

Tomando como referencia a Loría (2007) se tiene la representación general de un modelo VAR(p)⁸ irrestricto:

$$Z_t = \sum_{i=1}^n A_i Z_{t-i} + \varepsilon_t \quad (11)$$

donde: $Z_t = \begin{bmatrix} y_t \\ x_t^* \end{bmatrix}$ es decir, es un vector donde se logra capturar la endogenidad.

⁶ El tratado de todas las variables como parte de un todo, a un mismo nivel, hace posible medir impactos e interacciones de cada una en el conjunto.

⁷ Arias (2004) hace mención que los modelos VAR generan pronósticos confiables en el corto plazo. Sin embargo, tienen limitaciones como; omiten la posibilidad de considerar relaciones no lineales entre las variables y no toma en cuenta problemas de heterocedasticidad condicional ni cambio estructural en los parámetros estimados.

⁸ Al igual que en el modelo ARIMA, el orden de rezagos se pone entre paréntesis.

Además, Z_t es un vector columna $k \times 1$, n es el orden del modelo VAR, o número de retardos de cada variable en cada ecuación, ε_t es un vector $k \times 1$ de innovaciones que cumple con esperanzas matemática constantes, varianzas no dependiente en el tiempo y finita y las covarianzas entre dos variables aleatorias del proceso correspondiente a periodos distintos de tiempo solamente dependen del lapso de tiempo transcurrido entre ellas.⁹

El elemento y en la matriz A_i , $1 < t < n$ mide el efecto directo o parcial de un cambio unitario en Z_x en el instante t sobre Z_t al cabo de t períodos, $Z_{y,t+i}$. La columna x de la matriz A_i , mide el efecto que un cambio unitario en Z_x en el momento t tiene sobre el vector Z_{t+i} . De igual forma esta especificación considera que se cumple con el supuesto de ser estacionarias, al menos en el sentido débil (Novales, 2014).

Es decir, si se considera un modelo simple con dos variables y un rezago VAR (1), se tiene que:

$$\begin{aligned} y_{1t} &= \beta_{10} + \beta_{11}y_{1t-1} + \beta_{12}y_{2t-1} + u_{1t} \\ y_{2t} &= \beta_{20} + \beta_{21}y_{1t-1} + \beta_{22}y_{2t-1} + u_{2t} \end{aligned} \quad (12)$$

de forma matricial:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (13)$$

donde los términos de error satisfacen las propiedades antes mencionadas, resultando como forma matricial:

$$y_t = A_0 + A_1 y_t + \mu_t \quad (14)$$

⁹ Se asume que están contemporáneamente correlacionados, pero no autocorrelacionados, por lo que no tienen una matriz de covarianzas diagonal.

Por tanto, un shock en y_{2t} proveniente de u_{2t} , afectaran tanto a la propia y_{2t} como a los valores futuros de y_1 , esto debido a la integración de y_{2t-1} como variable explicativa de la ecuación de y_{1t} . Es importante mencionar que una condición fundamental de esta metodología es que los términos de error tengan una estructura de ruido blanco (Novales, 2014). Este aspecto es un punto fundamental, ya que la combinación lineal de todas las variables debe de ser estacionaria,¹⁰ de lo contrario deberán aplicarse primeras diferencias o tasas de crecimiento a las series originales (Loría, 2007).

Retomando la estructura del VAR descrita, en este modelo con dos variables, se tendrá una constante, más un rezago de cada una de las variables del modelo, lo que dará como resultado 6 coeficientes en total. Por lo que se debe de considerar que se tendrá tantas ecuaciones como variables, y los valores retardados de todas las ecuaciones aparecen como variables explicativas en todas las ecuaciones. Esto puede traer problemas ya que a medida que se incremente el número de variables y de rezagos, tenderá a incrementarse el coeficiente de determinación (R^2), pero al mismo tiempo podrán aparecer problemas de colinealidad, con lo cual generará parámetros no consistentes y habrá pérdida de grados de libertad (Loría, 2007).

La interpretación de los parámetros obtenidos es otro punto diferente, respecto a otras metodologías, ya que en estos modelos es común que algunos parámetros carezcan de significancia o tenga signos que van contra a la teoría planteada. Por lo que retomando la interpretación de Loría (2007) la lectura de estos no es directa, lo que realmente importa es su interpretación en conjunto.

1.3 Filtro de Kalman

Dentro de los modelos econométricos tradicionales un supuesto crucial en su construcción es la consideración de que los parámetros son constantes en el tiempo. Sin embargo, existen

¹⁰ El autor sugiere que todas ellas tengan el mismo orden de integración.

diversas razones para considerar que en la econometría actual los parámetros de muchos modelos no son constantes sino que varían con el tiempo (Maddala, 1977). Es decir, en condiciones reales la aleatoriedad es un factor muy importante que un sistema determinista no puede captar o analizar.

Dentro de un sistema determinista se asume que se posee un conocimiento perfecto de todas las cantidades necesarias para describir un sistema. Sin embargo, la realidad exige analizar los eventos como procesos estocásticos donde la aleatoriedad influye en el camino de las variables económicas. En este aspecto, el filtrado de variables puede ayudar a desarrollar algoritmos capaces de estimar una cierta variable (o vector de variables) que se ve interferida por ruidos y errores, los cuales queremos aislar para obtener un estimador lo más limpio posible (Raga, 2018).

Dentro de las dificultades para medir efectos no determinísticos, podemos agregar el problema de tener variables no observables lo cual es un elemento común dentro de la economía, pero también representan una tarea complicada para la construcción y estimación de estas. Se llama variables no observables a aquellas variables creadas como concepciones teóricas más que con una cuantificación económica y las cuales pueden construirse en función de sus efectos sobre variables observables (Coates, 2008) pudiendo servir como variables de apoyo entre construcciones teóricas y la realidad. Kalman (1960 y 1963) muestra un método conveniente para extraer señales, el cual consta de escribir las variables conocidas (observadas) y desconocidas en una misma representación en espacios de estado.

Rudolf Kalman (1960), realiza el primer acercamiento a un filtro que, por medio de datos medios, logre la estimación de una variable. El cual se basa en un proceso de dos etapas, la predicción del estado del sistema y la incorporación de las observaciones estimadas¹¹ teniendo como objetivo encontrar el estimador óptimo en términos del Error Cuadrático

¹¹ Una vez reduciendo el error de predicción

Medio (ECM), apoyado en el conocimiento del sistema previo del sistema y de las observaciones, junto con la descripción de su ruido y errores.

Autores como Andrews (2001), hacen mención de las propiedades óptimas de los estimadores generados por esta metodología.¹² argumentando que una ventaja de este método es la capacidad para elegir una correcta especificación cuando se conocen las varianzas de los ruidos que afectan al sistema. Además de la posibilidad de encontrar el mejor estimador lineal insesgado (MELI) cuando las variables tienen procesos estocásticos que se distribuyen como una normal y los vectores de error tienen impactos aleatorios (Deossa, 2009).

De tal forma la metodología del filtro de Kalman (Kalman, 1963) permite la estimación de modelos que poseen perturbaciones aleatorias, teniendo como premisa que el valor de parámetros β_t , que es potencialmente distinto para cada periodo t , y está vinculado al vector de parámetros de periodos anteriores a través de una matriz de transición que se supone *a priori*.

La forma para conseguir el mejor parámetro es a través de la función de máxima verosimilitud (Jalles, 2009), la cual es una técnica para estimar los valores de la varianza dada una muestra finita de datos, procurando encontrar los valores más probables de los parámetros de la distribución de un conjunto de datos (Arias, 1992). Esta contiene a los parámetros desconocidos, y aquel grupo de los mismos que maximicen la función de máxima verosimilitud, serán los utilizados (Paganini, 2003). Teniendo como ventajas: a) es una aproximación consistente para problemas de ajuste de parámetros, lo que significa que puede ser utilizada para una gran variedad de situaciones y b) provee estimadores insesgados y de varianza mínimas si se tiene una serie de datos suficientemente grande.

De igual forma, para que los parámetros sean consistentes estadísticamente y teóricamente, un elemento primordial de esta econometría es que las variables tengan el mismo orden de

¹² Respecto a cualquier función cuadrática de estimación del error.

integración.¹³ Debido a que muchas series temporales, en particular las económicas, no satisfacen las condiciones de estacionariedad ni en sentido fuerte ni débil, estas deben de ser expresadas en diferencias, esto buscando que su media y varianza sean constantes en el tiempo, eliminando el problema de regresiones espurias.

Es decir, tomando como referencia a Munera (2018), se puede afirmar que el filtro de Kalman nos dará un estimador que minimizará el error cuadrático medio a partir de unas observaciones:

$$\widehat{X}_{k|j} = \min_{\widehat{X}_{k|j} \in \mathbb{R}^n} \text{traza} \left\{ E \left[(X_k - \widehat{X}_{k|j})(X_k - \widehat{X}_{k|j})^t \right] \right\} \quad (15)$$

en donde si se considera un vector aleatorio X de dimensión l y siendo $\widehat{X}_{k|j}$ el estimador de X en el instante k , t indica la transposición. Así que el estimador a obtener es el valor esperado del estado en el tiempo k , condicionado por las observaciones. Por tanto, se puede reescribir como:

$$\widehat{X}_{k|j} = E[X_k | Z^j] \quad (16)$$

donde la variable Z^j se refiere al conjunto de observaciones. De tal forma, la matriz de varianzas-covarianzas del error queda:

$$P_{k|j} = E \left[(X_k - \widehat{X}_{k|j})(X_k - \widehat{X}_{k|j})^t \right] \quad (17)$$

Metodología del filtro de Kalman

Una de las principales características en la elaboración del filtro de Kalman es la definición de la ecuación estado-espacio, esta ayuda a realizar formulaciones generales que contengan características del sistema estocástico que permitan la predicción, suavizado y extracción de

¹³ Véase concepto de estacionariedad débil en el apartado dedicado a los modelos ARIMA.

señales de las variables no observables a través de la estimación conjunta del sistema y de las variables de estado. (Harvey, 1989). La representación en estado-espacio viene dada por dos ecuaciones vectoriales,¹⁴ una llamada de transición y la otra de observaciones. La primera de éstas, describe la dinámica del vector de estados \mathcal{E}_t , el cual contiene las variables desconocidas que queremos estimar. La segunda, relaciona al vector de estados con las variables observadas y_t^+ .

En cuestión de pronósticos obtenemos que el filtro de Kalman es un método econométrico para predecir las variables endógenas y para ajustar el parámetro estimado en las ecuaciones de pronóstico (Simionescu, 2015). De acuerdo a Harvey (1989) el filtro de Kalman tendrá dos procesos importantes:

- a) Estimación de valores de variables endógenas utilizando la información disponible previa.
- b) Ajuste de parámetros estimados usando ecuaciones de ajuste y computando errores de predicción.

Esto indica la existencia de un sistema de dos ecuaciones: uno que se encarga de predicción de los parámetros y uno de actualizaciones de la información disponible. Dichas ecuaciones se nombran y se expresan de la siguiente forma:

Ecuación de medición (relación entre lo observado y variables no observadas):

$$\xi_{t+1} = F\xi_t + C'x_{t+1} + Rv_{t+1} \quad (18)$$

¹⁴ Se define a la ecuación vectorial de la recta como aquella que está definida a partir de un punto $P_0(x_0, y_0, z_0)$ de la recta r , un vector director $\vec{v}(v_x, v_y, v_z)$ y un punto genérico $P_i(x, y, z)$ de la propia recta. Definida por la fórmula $OP_1 = OP + \lambda \cdot \vec{v}$ o $(x, y, z) = (x_0, y_0, z_0) + \lambda \cdot \vec{v}(v_x, v_y, v_z)$.

Ecuación de transición (estado dinámico (inobservable)):

$$y_t^+ = A_t' x_t^* + H_t' \xi_t + N_t w_t \quad (19)$$

donde ξ_t es el vector estados de las variables desconocidas y y_t^+ es el vector de estados de las variables observadas. Las series relacionadas están representadas por x_t^* y x_t , las cuales son variables exógenas. Se integra un término de error que se representa con v_t y w_t .¹⁵ Bajo el supuesto de existencia de no autocorrelación¹⁶ $F_t, C_t', R_t, A_t', H_t', N_t$, son matrices que contienen a los estimadores obtenidos por el método de máxima verosimilitud.

Por tanto, la estimación de un filtro de Kalman exige la implementación de tres etapas (Harvey, 1989):

- a) La estimación: con la suposición *a priori* del momento inicial $0 = \beta_{0/0}$ y $P_{0/0}$
- b) La predicción : $y_{0/0}$ en el momento 1 usando $\beta_{1/0}$ *
- c) La actualización: calculando la predicción del error, usando los valores observados en el momento 1 $n_{1/10} = y_1 - y_{1/10}$

De tal forma que tomando (18) y (19) se realizarán cálculos en cada interacción tomando las tres etapas antes mencionadas, creando un proceso de aprendizaje donde existe la actualización y predicción para el tiempo t ¹⁷ (Paganini, 2003). Dicho proceso se conoce como ganancia de Kalman y se define como el algoritmo optimo que minimiza el error cuadrático medio,¹⁸ lo cual hará que se pondere más la observación o el estimador según sea la variable con menor error. Debido a esto la ganancia de Kalman será un valor entre 0 y 1 que multiplicará la diferencia entre la observación y el estimador (Munuera, 2018).

¹⁵ Se asume que la distribución es normal multivariada (Paganini, 2003)

¹⁶ $\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & G \end{pmatrix} \right)$

¹⁷ El autor asume conocidos los valores para el tiempo t de $y_0^+, y_1^+ \dots y_t^+$

¹⁸ El ECM mide la diferencia entre el valor real y el valor estimado.

Considerando que x_t^* y x_t son conocidas y son series relacionadas¹⁹ hasta el tiempo $t + 1$ y las predicciones hasta $t - 1$: $\hat{\xi}_{t|t-1}, \hat{y}_{t|t-1}^+$. Además del conocimiento de la matriz de errores $P_{t|t-1}$. (Paganini, 2003).

Por lo que el paso de actualización quedará definido por:

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + P_{t|t-1} h_t (h_t' P_{t|t-1} h_t)^{-1} (y_t^+ - \hat{y}_{t|t-1}^+) \quad (20)$$

$$\hat{P}_{t|t} = \hat{P}_{t|t-1} + P_{t|t-1} h_t (h_t' P_{t|t-1} h_t)^{-1} \hat{P}_{t|t-1}^+ \quad (21)$$

donde la ganancia de Kalman (es decir, el valor ponderado y agregado a la nueva información obtenida) queda definido por $P_{t|t-1} h_t (h_t' P_{t|t-1} h_t)^{-1}$.

Dada esta nueva información incorporada al modelo se realiza un proceso de proyección de los valores esperados. De tal forma que tomando los conceptos definidos en las ecuaciones (18) a (21), se obtiene una nueva proyección:

$$\hat{\xi}_{t+1|t} = F \hat{\xi}_{t|t} + C' x_t \quad (22)$$

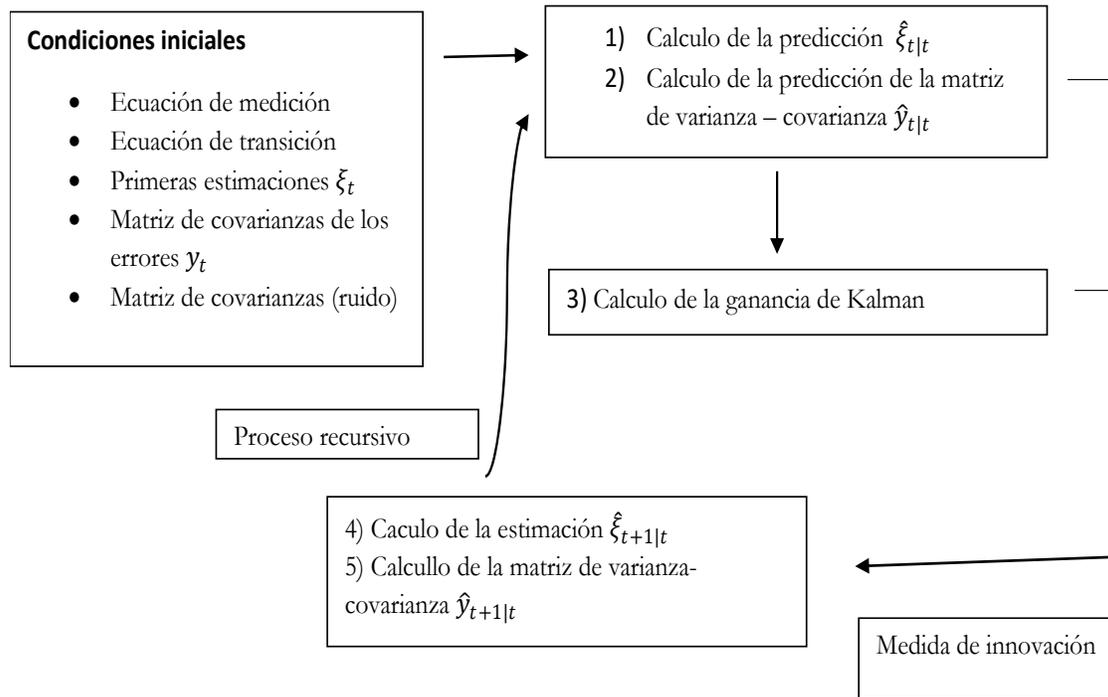
$$\hat{y}_{t+1|t} = \alpha'_{t+1} x_{t+1}^* + h'_{t+1} \hat{\xi}_{t+1}^* \quad (23)$$

$$P_{t+1|t} = F P_{t|t} F' + R Q R' \quad (24)$$

Con el apoyo del desarrollo antes mencionado, de forma intuitiva (diagrama 1) se puede inferir el proceso que el filtro de Kalman sigue para obtener estimadores eficientes (Raga, 2018).

¹⁹ Existe relación estadística entre variables explicativas al tener correlación respecto a la explicada.

Diagrama 1. Algoritmo del Filtro del Kalman



Fuente: Elaboración propia tomando como referencia a Raga (2018).

Para la estimación del primer momento (a) se considera un modelo con series relacionadas y estructura autoregresiva (Paganini, 2003), lo que ocasiona la suposición de que la serie a explicar sigue un proceso autoregresivo, ocupando correcciones para series no estacionarias en niveles suponiendo que y_t sigue un proceso AR(1). Siendo la ecuación estado-espacio igual a la antes mencionada (18).

Cabe mencionar que se considera como modelo autoregresivo aquel en el cual la variable explicada guarda una relación lineal con sus propios valores anteriores (Munuera, 2018). Es decir:

$$X_k = \alpha + \sum_{i=1}^p \beta X_{k-i} \quad (25)$$

donde X será la variable a explicar y que p será el orden del modelo.

Con esta información podemos describir las ecuaciones (17) y (18) para establecerlas de una forma más intuitiva y para una mejor lectura se tiene que las ecuaciones a ocupar son:

$$X_{t+1} = K_t X_t + C_t U_t + w_{t+1} \quad (26)$$

$$Z_t = H_t X_t + A_t U_t' \quad (27)$$

Por tanto se tiene que X es la variable a explicar en el modelo, las variables U_t y U_t' se refieren a las series relacionadas, el término w_t es el termino de error que sigue una distribución normal de media 0 y no correlacionado. Además que K_t, C_t, H_t, A_t serán las coeficientes de las matrices a estimar y que maximizaran la densidad por medio de la función de máxima verosimilitud.

2. Hechos Estilizados

Resulta muy interesante realizar un análisis de las regularidades empíricas en las series a estudiar ya que se pueden encontrar similitudes o comportamientos definidos entre las variables que ayuden a realizar un razonamiento teórico. Un correcto análisis gráfico permite estudiar visualmente la evolución de las variables consideradas en el período seleccionado, ayudando a detectar comportamientos cíclicos, tendenciales, estacionales o, en muchas ocasiones, puntos atípicos (Mahía, 2001).

Dado la revisión teórica y la experiencia empírica se ha seleccionado a las siguientes variables para la realización de este modelo: a) la tasa de desempleo²⁰ (deso) la cual será la variable pronosticada dentro de este ejercicio econométrico (variable explicada), b) las Remuneraciones Medias Reales por Persona Ocupada (salreal) la cual funge como variable proxy del salario real y c) el Indicador Global de Actividad Económica (igae) esta variable

²⁰ Se remplace el nombre de “Tasa de población desocupada” por el de “Tasa de desempleo” debido a la estandarización de términos respecto a otras publicaciones y autores.

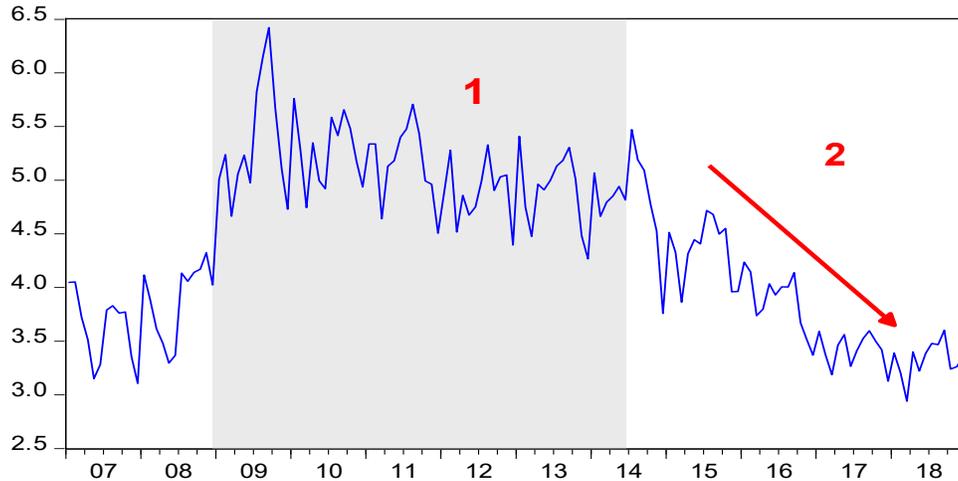
se ocupa como el proxy del PIB dado que la periodicidad de esta última es de frecuencia más baja. Todas las variables fueron consultadas en INEGI y Banxico²¹ sin ninguna transformación previa²² esto con la intención de que recreen de mejor forma el comportamiento natural del fenómeno que están representando e intentar realizar pronósticos más acertados acordes a la dinámica intrínseca de las series (Diebold, 2004). De la misma forma, estas variables se encuentran con periodicidad mensual con el fin de captar por separado cada uno de los efectos estacionales que pudieran contener. La muestra es la máxima posible acorde a lo reportado, así que van desde enero de 2007 hasta marzo de 2019, por lo que se tiene una muestra de 171 datos lo cual pasa ampliamente los límites mínimos para validar pronósticos econométricos (Diebold, 2004).

Un primer acercamiento a conocer el desempeño de la variable a pronosticar es su análisis gráfico y su análisis estacional, esto con el fin de lograr asimilar y notar patrones intrínsecos antes de realizar estimaciones econométricas. En la gráfica 1 se puede apreciar que la tasa de desempleo ha tenido dos periodos muy marcados a partir de 2007; en un primer momento (1) se logra apreciar como a partir de la crisis de 2009 la población desocupada tocó máximos históricos dentro del intervalo de esta serie. A partir del año 2014 (2) esta comienza a bajar de forma gradual hasta tocar y caer aún más que en el periodo pre crisis. Se puede apreciar que existe un intervalo bien definido entre los valores reportados en esta serie, por lo cual podemos obtener un valor mínimo o máximo.

²¹ La fecha de consulta para las tres series fue en mayo de 2019.

²² Con efectos de un correcto análisis exploratorio previo las variables se ocupan en niveles. Sin embargo, para la realización del filtro de Kalman se ocupan sus primeras diferencias con el fin de maximizar las propiedades de esta técnica y evitar problemas de regresión espuria.

Gráfica 1. Tasa de Desempleo (2007m01 - 2019m03)

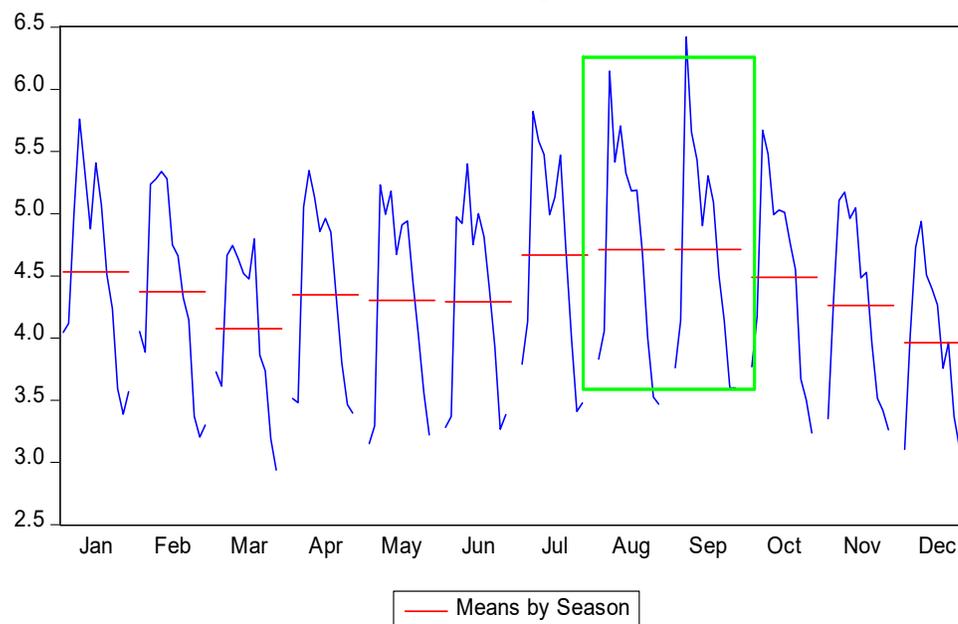


Fuente: Elaboración propia con datos de INEGI (2019)

Otro aspecto muy importante en la elaboración de pronósticos es la gráfica que muestra la estacionalidad, la cual nos ayudará a notar patrones característicos en los diferentes meses que componen el año (Gráfica 2) esto debido a que un correcto análisis nos puede ayudar a encontrar comportamientos recurrentes en el corto plazo de la serie pronosticada, por tanto, se pueden notar escenarios donde el valor de la serie puede ser mayor o menos según sea el efecto estacional sobre esta. Para este caso se puede apreciar que tenemos estacionalidad en el mes de agosto y septiembre, este efecto se puede deber a que las empresas planifican a futuro, por lo cual el último trimestre del año es fundamental para realizar recortes de personal y poder tasar aumentos o disminuciones en la producción, de tal forma se puede apreciar un incremento en los valores de la serie.²³

²³ Este comportamiento se toma en cuenta en la elaboración del estado-espacio y la retroalimentación en el modelo.

Gráfica 2. Estacionalidad de la Tasa de Desempleo (2007m01 - 2019m03)



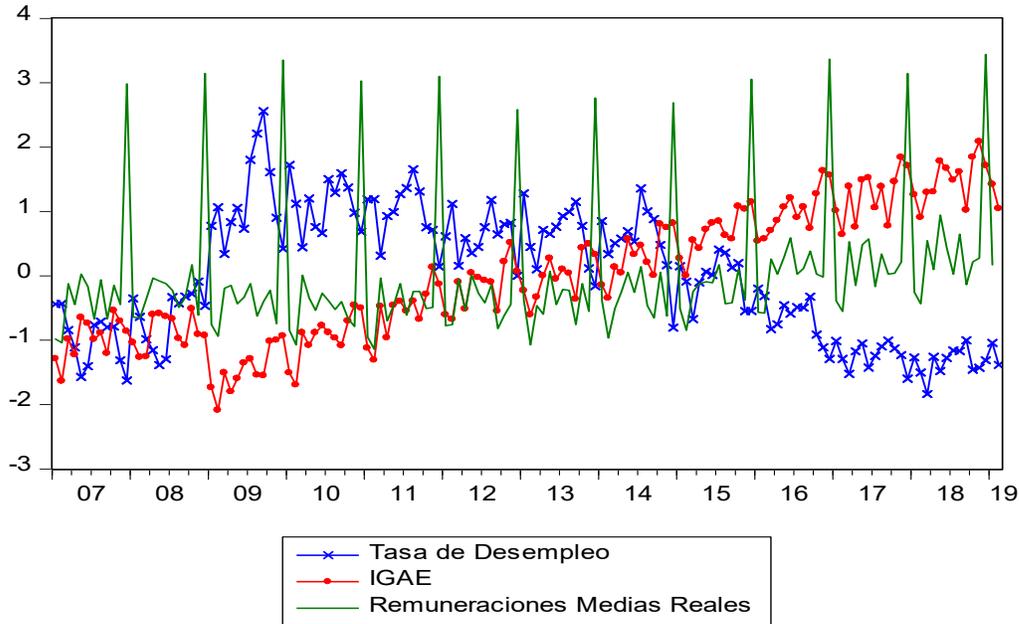
Fuente: Elaboración propia con datos de INEGI (2019)

De la misma forma es necesario conocer la relación estadística²⁴ que guardan las variables seleccionadas, por tal motivo se requiere graficarlas en conjunto y conocer la interrelación visual que existe entre ellas, esto nos ayudará a comprobar (de forma visual) si el comportamiento de las variables explicativas pueden ayudar a la predicción de la variable a pronosticar. La gráfica 3 nos muestra la existencia de una relación negativa entre las variables explicativas (igae y salreal) respecto a la variable a pronosticar (deso). Este efecto es mencionado por Solimano (1986) el cual indica que en una economía abierta, la relación entre desempleo y salario real consiste en ser el punto de transferencia para la eliminación de cuellos de botella de la expansión económica, esto debido a que es imprescindible aumentar el nivel de competitividad internacional, lo cual se logra abaratando relativamente los costos de producción local, lo cual implica un aumento en la productividad y/o reducir el precio de los factores productivos locales. Es decir, para aumentar el nivel de empleo en una economía abierta que tiene restricción dominante a

²⁴ Se define como relación estadística cuando dos o más variables están relacionadas estadísticamente, por tanto la primera se puede estimar aproximadamente el valor de las otras (Spanos, 2002).

corto plazo es necesario la reducción del salario real. Este efecto se hace muy notorio a partir de 2014, coincidiendo con el periodo II encontrado en la gráfica 1.

Gráfica 3. Variables normalizadas de la Tasa de Tasa de Desempleo, IGAE y Remuneraciones Medias Reales (2007m01 - 2019m03)



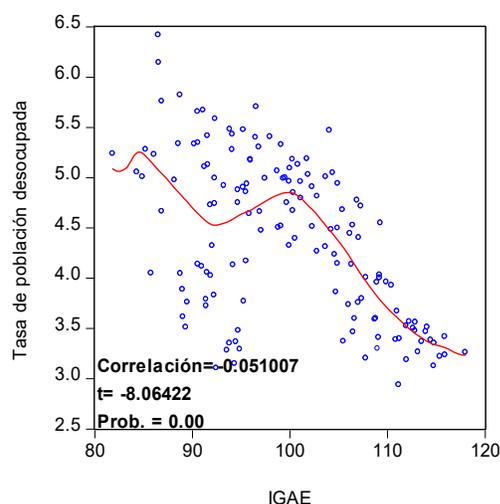
Fuente: Elaboración propia con datos de INEGI y Banxico (2019)

Para confirmar la relación existente, el signo de la correlación y la magnitud en la que las variables seleccionadas impactan sobre la variable pronosticada es necesario realizar un análisis de correlación (cuadro 1), de esta forma se aprecia que el análisis gráfico coincide con el análisis estadístico y la magnitud de las variables seleccionadas es estadísticamente significativa.

Cuadro 1. Correlación entre la Tasa de desempleo respecto al IGAE y Salario real para México (20017m01 - 2019m03)

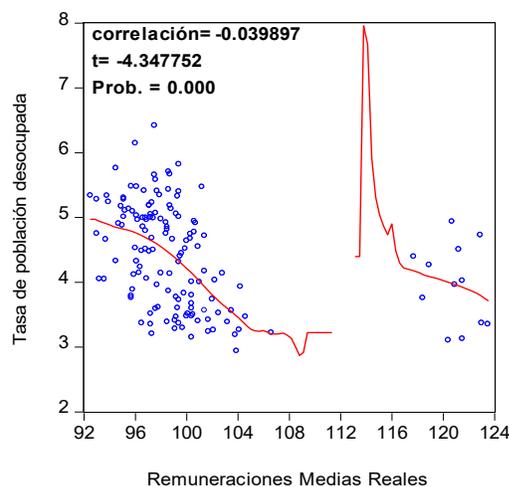
	Tasa de desempleo	IGAE	Salario Real
Tasa de desempleo	1		
IGAE	-0.553521	1	
Salario Real	-0.341694	0.340256	1

Gráfica 4. Tasa de desempleo e IGAE (Ajuste de Kernel y en niveles)



Fuente: Elaboración propia con datos de INEGI (2019)

Gráfica 5. Tasa de desempleo y Remuneraciones Medias Reales (Ajuste de Kernel y en niveles)



Fuente: Elaboración propia con datos de Banxico (2019)

Un efecto muy interesante dentro del análisis de hechos estilizados es la comprobación de una ausencia de linealidad cuando las variables se encuentran en niveles²⁵ (gráfica 4 y 5). Esto hecho comprueba la necesidad de trabajar con las diferencias de las variables, lo cual concuerda con la estructura planteada capítulos anteriores. La existencia de no linealidades en niveles en relación empírica de las variables puede ser un indicador para buscar modelos alternativos a los métodos convencionales y apoyaría la hipótesis sobre probar nuevas metodologías econométricas que logren maximizar las propiedades de la serie.

Conforme al análisis visual y estadístico realizado se puede comenzar a analizar primeras conclusiones que dan pie a la posibilidad de implementar el filtro de Kalman en el fenómeno económico analizado. En primer momento se encuentra que la tasa de desocupación tiene un cambio en la dinámica, lo cual podría ser mejor captado por un algoritmo de autoaprendizaje como lo es la ganancia de Kalman. De igual forma, se analiza que el modelo econométrico debe de ser planteado en un inicio con variables relacionadas y proceso autorregresivo, además que se cuenta con series no estacionarias en niveles. Estos

²⁵ Ante la presencia de valores en niveles se recomienda realizar primeras diferencias para evitar regresiones espurias.

tres aspectos son fundamentales en el planteamiento de una ecuación estado-espacio (Paganini, 2003) y serán tomados en cuenta en procesos posteriores.

3. Aspectos Econométricos

Bajo las premisas teóricas (económicas y econométricas) planteadas en apartados anteriores se procede a estimar las tres metodologías, esto con el fin de realizar una correcta aproximación econométrica que ayude a comprobar la eficiencia de estas.²⁶ Primeramente, se estimará un modelo univariante (ARIMA), seguido se realizará dos metodologías multivariantes (VAR y filtro de Kalman).

Todos los modelos se realizaron en el programa Eviews 10, el forecast se realizó por medio del método estático, esto debido a que la predicción estática es una predicción paso a paso que utiliza los valores verdaderos de las variables rezagadas para efectuar el pronóstico (Diebold, 2004). Se ocupó una técnica de pronósticos combinados, recomendada por EuroStat (2016) y es una forma útil de incorporar pronósticos basados en diferentes modelos en un compuesto pronóstico que supera a los pronósticos individuales. De esta forma se evaluó mes a mes cada pronóstico generado, comparándolo con tasas de variación mensuales y anuales, realizando la valoración de cada uno de los componentes dentro del pronóstico, tomando en consideración una menor variación la raíz del promedio de los errores al cuadrado, el cual mide la desviación estándar de los errores y el Coeficiente de Desigualdad de Theil, el cual mide la diferencia entre el pronóstico y los valores reales. Si el coeficiente tiende a cero, el modelo puede ser utilizado para predecir dado que sus pronósticos serán fiables, por el contrario, si el coeficiente tiende a uno el modelo no sirve para predecir.

²⁶ Los modelos fueron estimados en mayo de 2019, de tal forma, que para la aprobación de este trabajo se ha tenido un año para comprobar la eficiencia de los pronósticos arrojados.

3.1 Modelo ARIMA

Se especificó y probó una serie de modelos ARIMA, tanto de la forma automática como manual, optando por la opción de un modelo ARIMA por medio de mínimos cuadrados condicionales (CLS) y con la opción automática, ya que las opciones predeterminadas utilizadas por los programas estadísticos generalmente producen resultados satisfactorios, esto debido a que en presencia de series altamente estacionales y con tendencia esta proporciona valores cercanos a los observados²⁷ (FMI, 2017).

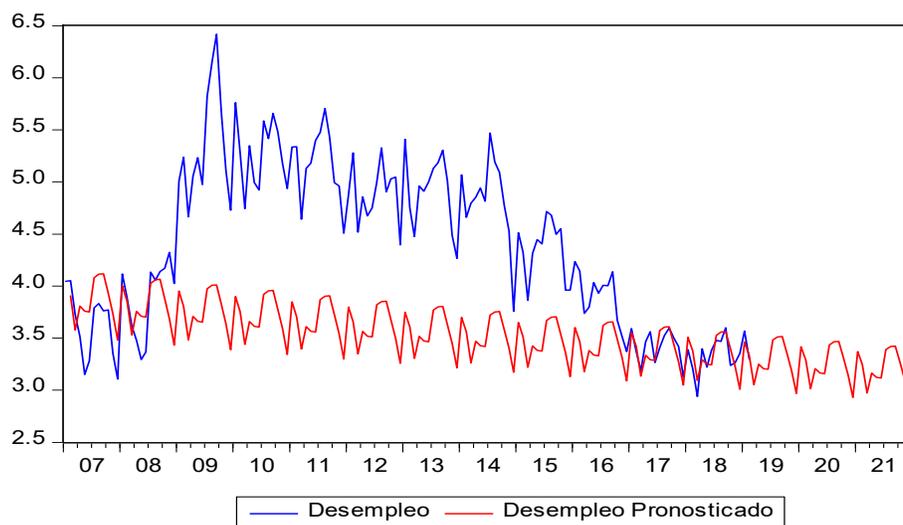
El algoritmo de CLS está basado en la minimización la suma de los errores cuadrados con una función estimada que contiene un término de error el cual es una versión del modelo ARIMA (Arias, 1992). Considera primero una media móvil integrada.

Se comprueba la raíz unitaria de las variables, lo cual se verifica mediante las pruebas de raíz unitaria Augmented Dickey-Fuller (ADF), Phillips-Perron, (PP), y Kwiatkowski-Phillips-Schmidt-Shin (KPSS), concluyendo que estas son estacionarias en primeras diferencias. El modelo es hecho estacionario e implica que Y_t fue integrada con diferenciación, proporcionando estimadores más eficientes y comparables a través del tiempo (Genschel, 2010). De la misma forma se elige el método de “Eviews Legacy”, el cual es una implementación particular de Gauss-Newton con Marquardt y proporciona compatibilidad de la estimación hacia periodos anteriores (Eviews, 2016).

Por lo cual el resultado es un modelo ARIMA (2,1,1) el cual pasa todas las pruebas de correcta especificación y presenta estabilidad dinámica en sus componentes.

²⁷ Sin embargo, este mismo organismo recomienda su uso solamente en uno o dos periodos.

Gráfica 6. Desempleo vs Desempleo pronosticado para México (ARIMA)



Fuente: Elaboración propia (2019)

Como se puede apreciar en la gráfica 6, el modelo ARIMA no predice de forma correcta la dinámica de la serie, la cual reproduce solo con patrones estacionales. Aunado al análisis gráfico el coeficiente de Theil es de 0.13, lo cual es alejado del 0.

3.2 Modelo VAR

Para continuar con el análisis se procede a realizar un modelo VAR, el cual quedará especificado con el siguiente sistema de información (Z):

$$Z = \{deso_t, salreal_t, igae_t\}$$

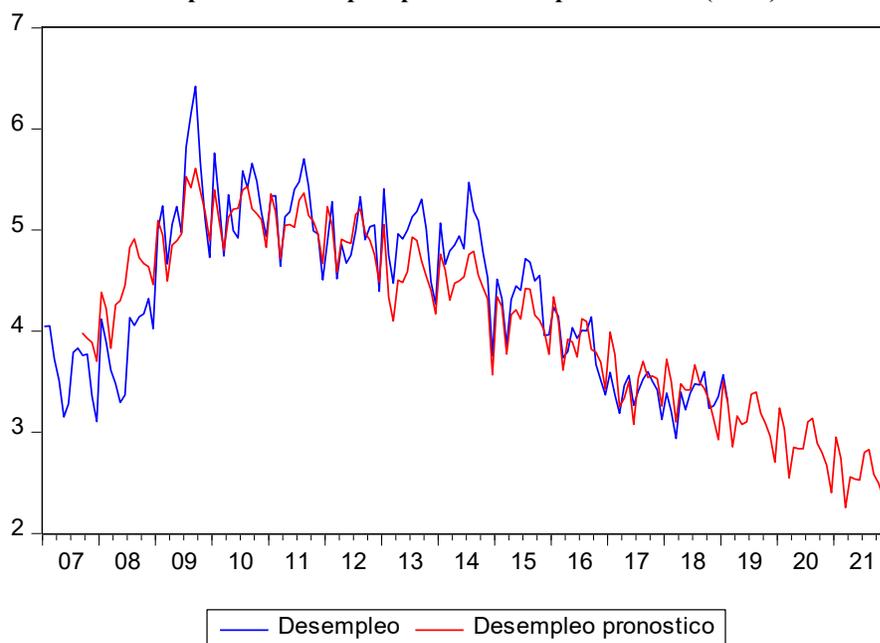
donde *deso* corresponde a la tasa de desempleo (variable a pronosticar), *salreal* corresponde al salario real e *igae* es el Indicador Global de Actividad Económica.

De igual forma que en la metodología ARIMA, las series son diferenciadas y el número de rezagos está determinado por los criterios LR, Final prediction error, Akaike, Schwarz y

Hannan-Quinn, los cuales indican que la especificación más eficiente se alcanza incluyendo ocho rezagos.

El modelo cumple con todas las pruebas de correcta especificación hechas al 5% de significancia, tales como: Módulo máximo=0.92; Urzúa=7.47 (0.27); LM(8)=11.69 (0.23); White(c)=301.02(0.91).

Gráfica 7. Desempleo vs Desempleo pronosticado para México (VAR)



Fuente: Elaboración propia (2019)

La gráfica 7 permite apreciar el pronóstico realizado por la metodología VAR,²⁸ el cual replica de mejor forma el comportamiento *in-sample* de la serie, además de tener un mejor coeficiente de Theil (0.03). Sin embargo, su capacidad para pronosticar los datos futuros difiere de la dinámica intrínseca de la serie, esto debido a que prevé una caída de la tasa de desempleo, lo cual en el contexto actual se vislumbra complicado.

²⁸ El pronóstico fue transformado a valores nominales debido a que el modelo lo arroja en primeras diferencias

3.3 Filtro de Kalman

Finalmente se procede a estimar el filtro de Kalman, esperando que este genere predicciones óptimas y actualizaciones de las variables de estado con el fin de generar un proceso de retroalimentación en el modelo (Simionescu, 2015). Para la elaboración del modelo se consideró su especificación por dos etapas:²⁹

- a) Estimación de valores de variables endógenas utilizando la información disponible previa.

El filtro de Kalman determina el estimador del error cuadrado mínimo del vector de variables de estado; por tanto, la literatura ha definido dos enfoques para la estimación de una variable utilizando este filtro. El primero asume que el valor inicial de la variable de estado no estacionaria puede ser fija y desconocida. Mientras que el segundo enfoque considera que el valor inicial es aleatorio.

Para conocer que enfoque será el mejor para estimar el modelo, se considera que la metodología del filtro (Kalman, 1960). Para la elaboración de valores previos (ecuación de estado-espacio), la cual será eje de inicio para el proceso iterativo, se procede a identificar componentes estacionales. De igual forma, se establece un componente cíclico a través de la modelación de los residuales con el algoritmo de selección del menor criterio de información de Schwarz³⁰ entre un grupo de regresiones. Los resultados indican que se debe de iniciar la modelación con un proceso autorregresivo de medias móviles de cuarto orden, ARIMA (4,1,4).

El lograr analizar las primeras observaciones del estado-espacio definido, puede generar resultados más eficientes (Koopman, 1999) además que reducen problemas posteriores debido a no conocer los valores de las varianzas. Por ende es necesario calcular los

²⁹ Esta serie de pasos fue descrita en capítulos anteriores.

³⁰ El resultado es que el BIC generalmente selecciona el modelo más abstracto y más sencillo (Diebold, 2004).

parámetros iniciales, esto buscando que los choques provenientes de σ_w^2 sean mayores que los choques de σ_t^2 (Simionescu, 2015). De tal forma se elaboró una regresión alterna para conocer el valor que asumen estos parámetros en su estado inicial. Este modelo fue elaborado por la metodología Función Transferencia³¹ (Novales, 1993) y posee todas las pruebas de correcta especificación: $R^2 = .9006$ JB = 1.6304(0.4425), LM (0.5440), White = (0.8529). El cual está definido por múltiples inputs siguiendo la forma propuesta por Bekkali (2013)

$$Y_t = v_1(B)X_{1,t} + v_2(B)X_{2,t}$$

donde el sistema de información queda definido por las mismas variables empleadas en el modelo VAR.

Con el modelo anterior se busca encontrar la especificación inicial que apoye en el proceso descrito. Por lo cual, se transforma un modelo unicuacional en un modelo de estado-espacio que logre capturar de forma eficiente las señales de las variables externas.

Por lo cual se asignó la siguiente especificación estado-espacio:

```
@param c(9) 1.059130 c(10) -0.158709 c(4) -0.644290 c(5) 0.385254

@signal ddeso = c(1) + c(2)*@seas(8) + c(3)*@seas(9) + sv1*digae2 + sv2*dsalreal2 + c(4)*sv3

@state sv1 = sv1(-1)
@state sv2 = c(9)*sv2(-1) + c(10)*sv3(-1) + c(11)*sv4(-1) + [var = exp(c(8))]
@state sv3 = sv2(-1)
@state sv4 = sv3(-1)

@mprior svec2
@vprior svar2
```

Se estimó el modelo de estado-espacio basándose en lo establecido por Bossche (2011) partiendo de una especificación de la matriz de covarianzas y el vector de estado igual a cero. Posteriormente se actualizaron ambos elementos con los valores correspondientes a sus estados finales y se incluyeron en la especificación del modelo para su respectiva

³¹ El modelo fue elaborado con base a los residuos de las variables seleccionadas para el modelo principal.

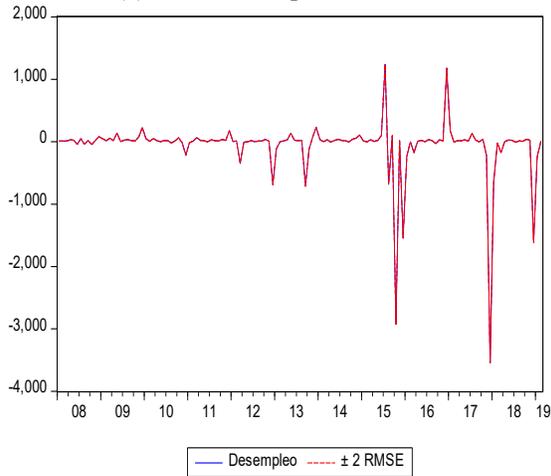
estimación recursiva, obteniendo 134 iteraciones y se eligió el modelo que muestra mejor principio de convergencia a un estado estable.

Todos los estados son significativos al 5% de significancia, incluidos los estados finales, si bien se pueden verificar estadísticos habituales a la serie generada por el filtro de Kalman (correlograma con las estadísticas de Box-Ljung y la normalidad de Jarque-Bera), estos diagnósticos se encuentran fuera del modelo. Esto implica que el correlograma, así como los grados de libertad asociados, son generalmente incorrectos (Bossche, 2011). De igual forma (por medio de la prueba de elipses) se puede comprobar que los valores del modelo probabilístico convergen a las observaciones.³²

La gráfica 8(a) muestra los datos la tasa de desempleo y las estimaciones suavizadas junto con bandas de confianza. De igual forma la gráfica 8(b) muestra los errores de predicción estandarizados, los cuales toman en consideración la relación de los residuos de predicción y los errores estándar de los residuos de predicción (Bossche, 2011). Se puede apreciar que ambas bandas están muy cercanas una de otra, lo cual da vista de que el modelo planteado tiene una buena predicción.

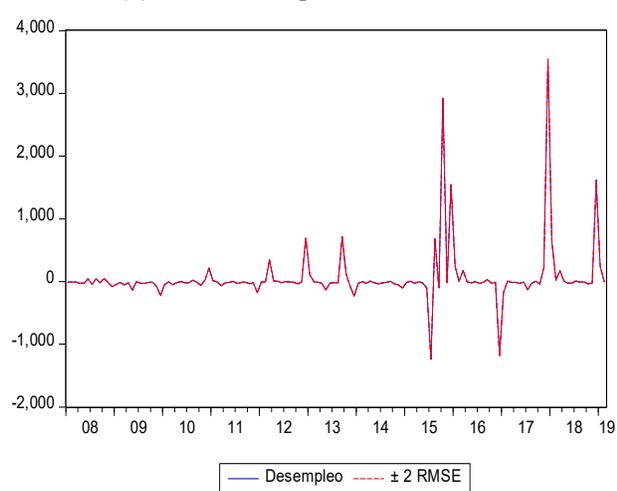
³² Mientras para algunos modelos la representación unidimensional genera la aceptación o rechazo de una hipótesis nula, en el caso de modelos con de restricciones, se puede formar regiones de confianza o elipses de confianza.

Gráfica 8(a). Señales de predicción



Fuente: elaboración propia

Gráfica 8(b). Errores de predicción estandarizados



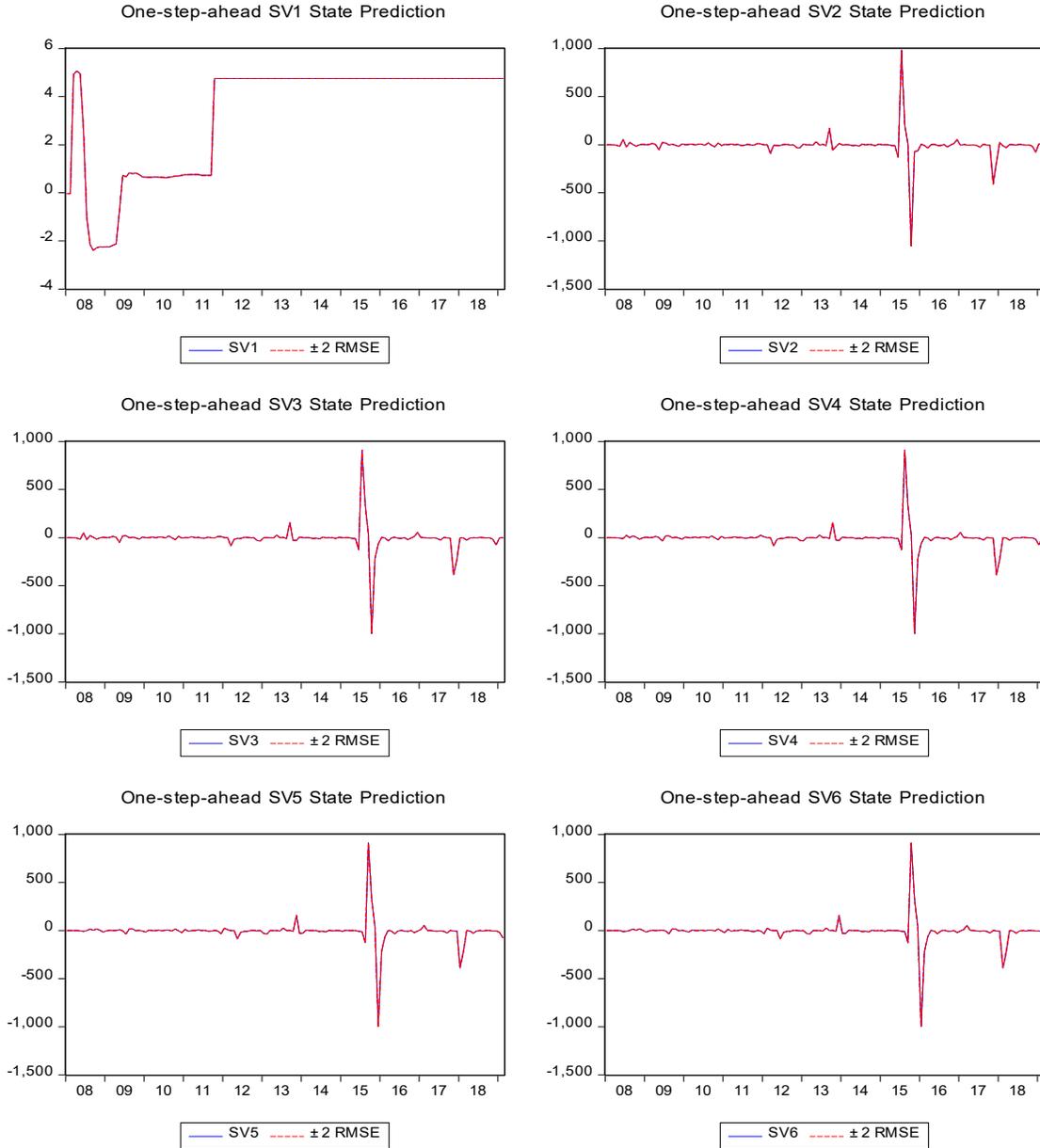
Fuente: elaboración propia

Una vez encontrado el modelo estado-espacio que maximiza la resolución de la ecuación por medio del algoritmo del filtro de Kalman, se tienen los parámetros que probabilísticamente replican de mejor forma el sistema de información elegido. De esta forma, se considera que tenemos la especificación necesaria para aislar las señales externas que podrían influir en los efectos a corto plazo de la variable a pronosticar, y las que generan el proceso de aprendizaje del filtro de Kalman.

Estas propiedades ayudarán a obtener pronósticos con mayor precisión y que estarán alimentados por los mismos errores que este tenga a lo largo del tiempo. Es decir, para la predicción del futuro se ocupa un proceso recursivo basado en el estado en tiempo t de la serie (posición, velocidad, aceleración, etc).

Por tanto, al tener series relacionadas se considera que estas influyen en el camino que toma la serie a pronosticar realizando una comparación de los valores estimados y reales en los tiempos t y t_1 , lo cual generará una diferencia que el algoritmo reducirá gradualmente hasta ajustar la estimación a los valores reales.

Gráfica 9. Estados de predicción



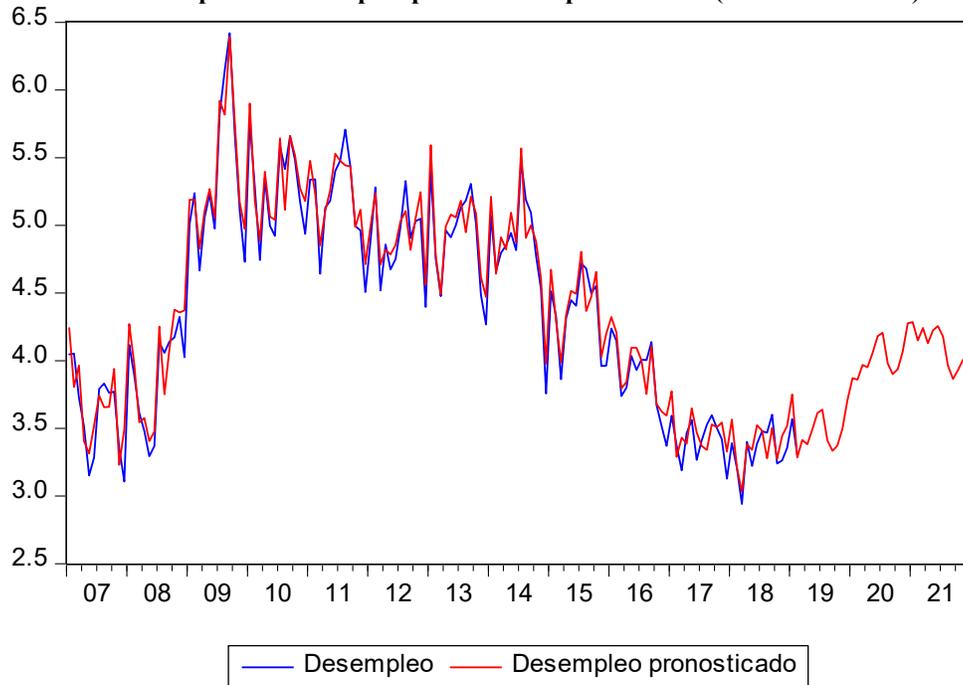
Fuente: elaboración propia

El proceso predictivo se comprobó con la gráfica 8 en la cual se puede apreciar que la generación de escenarios es muy cercana a los observados, esto debido a que ambas líneas tienen trayectos muy similares. En primer momento se muestra que la especificación de la ecuación estado-espacio asemeja a las señales observada. Las 3 gráficas restantes muestran el proceso de aprendizaje y como este con el pasar de los momentos de forma recursiva va

reduciendo la diferencia con la serie observada, lo cual como se mencionó con anterioridad es un punto fundamental para considerar al filtro de Kalman como una poderosa herramienta en la elaboración de pronósticos.

Cabe mencionar que conforme a las dos gráficas mostradas con anterioridad (8 y 9) se puede notar las capacidades predictivas del filtro de Kalman en este ejercicio. En primer momento se nota que el modelo lograr capturar de correcta forma el estado-espacio, además que las señales de las variables explicativas arrojan en el sistema recrean de buena forma a la variable a pronosticar, posteriormente se comprueba que el proceso recursivo sigue un patrón de aprendizaje que va reduciendo cada vez más la diferencia entre la estimación y las series observadas.

Gráfica 10. Desempleo vs Desempleo pronosticado para México (filtro de Kalman)



Fuente: elaboración propia

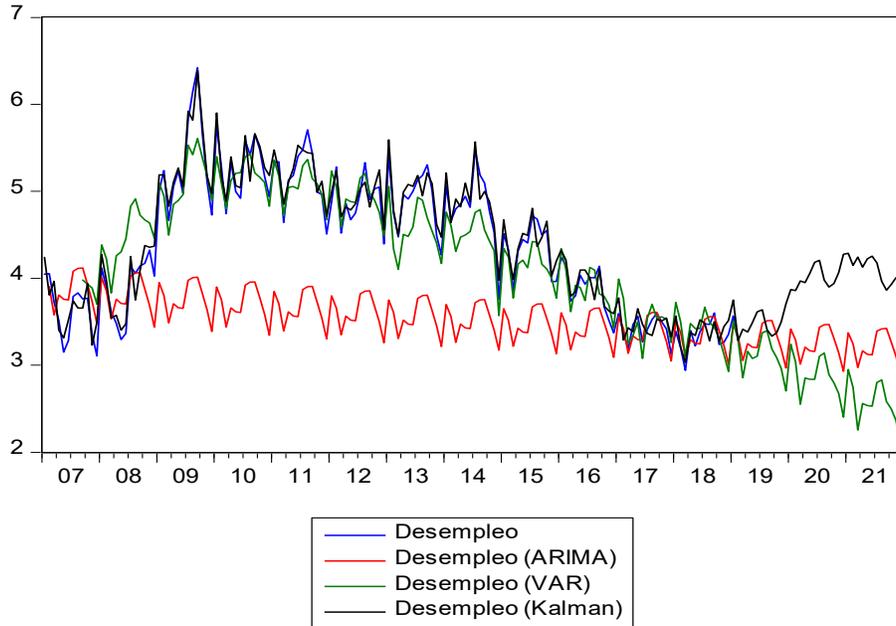
Como muestra la gráfica 10, se puede comprobar que el pronóstico hecho por el modelo replica de buena forma a los datos observados, obteniendo un coeficiente de Theil muy

cercano a cero (0.01). Además, que el pronóstico fuera de la muestra tiene una dinámica que puede obedecer al ciclo económico y el incremento de la población sin empleo.

3.4 Comparación de pronósticos

Al graficar los tres series estimadas y compararlas (gráfica 11) se observa que los comportamientos del modelo ARIMA, VAR Y del filtro del Kalman tienen diferencias sustanciales, esto es más notorio en el pronóstico fuera de la serie. Los modelos VAR y ARIMA arrojan un pronóstico que dista mucho de poder replicar el comportamiento visualmente similar al que la tasa de desocupación tiene en periodos anteriores, por lo cual el filtro de Kalman posee mejor desempeño dinámico de las observaciones hacia futuro.

Gráfica 11. Tasa de desempleo pronosticada por ARIMA, VAR y filtro de Kalman



Fuente: elaboración propia con datos de INEGI

Con base a las pruebas estadísticas y las visuales es pertinente tomar una decisión sobre el modelo que podrá servir como referente y que valores continuaran la serie. El filtro de Kalman desde su construcción presentó mejores parámetros de inicio y recursivos, de igual

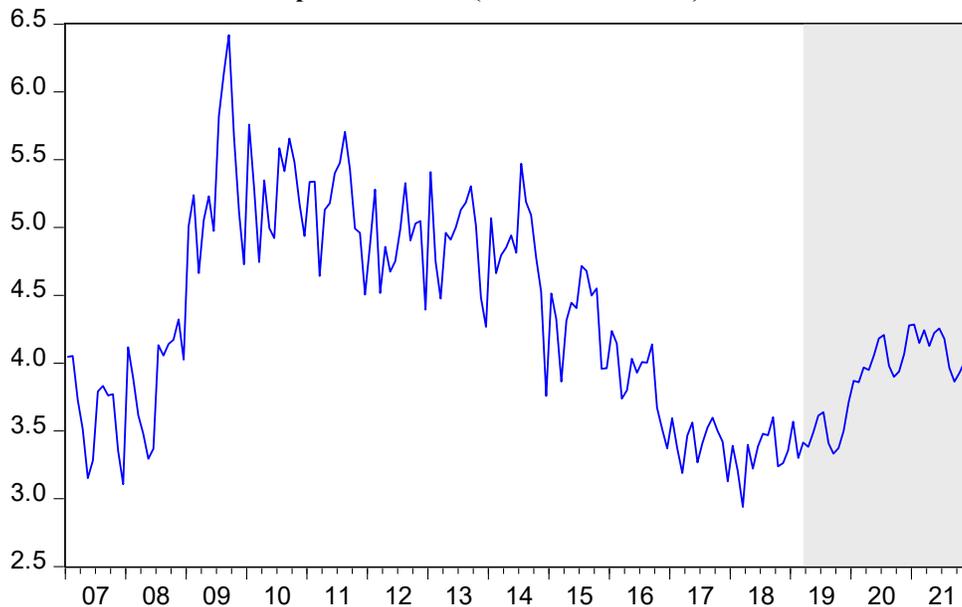
forma a nivel estadístico tuvo un coeficiente de Theil más cercano a 0 (cuadro 2) y, por último, gráficamente replicó de mejor forma la dinámica de la serie. Por tal motivo, se opta por tomar los valores de éste para considerar un pronóstico mensual de la tasa de desempleo (gráfica 12) por tanto el filtro de Kalman proporciona pronósticos más eficientes que las otras metodologías planteadas.

Cuadro 2. Comparación de coeficiente de Theil

Modelo	Coefficiente de Theil
ARIMA	0.13
VAR	0.03
Filtro de Kalman	0.01

Fuente: elaboración propia

Gráfica 12. Tasa de desempleo en México (2007m01-2021m12)



Fuente: elaboración propia

Desde la elaboración de este trabajo (mayo de 2019) hasta las revisiones de este (mayo de 2020) el pronóstico generado por el filtro de Kalman ha tenido una diferencia promedio de 0.08, lo cual es un margen de error muy pequeño en un pronóstico a largo plazo.

Conclusiones

En las secciones anteriores, se describió y aplicaron metodologías correspondientes a ARIMA, VAR Y Filtro de Kalman para pronosticar la tasa de desempleo en México hasta 2021. Comprobando que el filtro de Kalman arrojó mejores resultados en coeficiente de Theil (0.01) y replicó de mejor forma la dinámica de la serie.

Esta metodología logra obtener tasas muy cercanas a las observadas en la muestra, de igual forma tuvo variaciones de 0.08 a lo largo de un periodo de prueba comprendido de mayo de 2019 a mayo de 2020. Esta metodología otorga una libertad que logra capturar la dinámica intrínseca de la serie dado sin la necesidad de incluir variables ficticias que apoyen la correcta especificación, lo cual es una ventaja en la búsqueda de recrear movimientos futuros de las variables económicas.

Desde la realización de este trabajo (mayo 2019) el modelo ha presentado buenos pronósticos, sin embargo, el efecto coyuntural que ha traído la pandemia de COVID-19 ha ocasionado que los pronósticos después junio de 2020 tengan un desfase respecto a lo observado. Si bien no existe ninguna técnica capaz de predecir eventos que impacten de forma directa y tan fuerte a las variables económicas, se considera que el trabajo aquí realizado tuvo buenos resultados en periodos de “estabilidad”.

Si bien una desventaja de esta metodología, además de la poca o nula literatura de su aplicación en México, es la complejidad para la correcta definición de los estados-espacio iniciales y la cantidad necesaria y óptima para conseguir la optimización de estos, sus bondades abren un gran abanico de poderosas herramientas que se pueden implementar en los pronósticos económicos.

Como valor agregado a este trabajo se introdujo a las herramientas de análisis cuantitativo que van más allá de las convencionales, logrando una interacción mayor entre diversas ramas de estudio, y ponderando las ventajas y desventajas de su utilización en la elaboración de pronósticos. Aunado a esto, el campo de crecimiento para este trabajo es amplio, ya que en primera instancia se requiere una actualización periódica de los datos y su posterior análisis en los modelos planteados. De igual forma, la pandemia de COVID-19 abre la posibilidad de actualizar e integrar nuevas metodologías que ayuden a un pronóstico más certero.

Bibliografía

- Amarís, Á. y. (2017). *Aplicación de modelo ARIMA para el análisis de series de volúmenes anuales en el río Magdalena*. Revista tecnura 21(52), 88-101.
- Andrews, G. (2001). *Kalman Filtering Theory And Practice Using Matlab*. New York: Wiley.
- Arias, E. (1992). *Evaluación y maximización de la función de verosimilitud de procesos ARMA multivariantes*. Madrid: Universidad Complutense de Madrid.
- Arias, E. (2004). *Modelos VAR Y VECM para el pronóstico de corto plazo de las importaciones de Costa Rica*, Costa Rica: Banco Central de Costa Rica.
- Bekkali, N. (2013). *Modelos ARIMA y Función de Transferencia para previsiones temporales*. Universitat Politècnica de Valencia.
- Bernanke, M. G. (1997). *Systematic Monetary Policy and the Effects of Oil Price Shocks*. Brookings Papers on Economic Activity, 91-142.
- Blanchard, O. J. (1999). *Wage Dynamics: Reconciling Theory and Evidence*. NBER Working Paper Series.
- Bossche, F. V. (2011). *Fitting State Space Models with EViews*. Journal of Statistical Software, May 2011, Volume 41, Issue 8.
- Camba-Mendez, G. (2012). *Conditional forecasts on SVAR models using the Kalman filter*. Economics Letters, 376–378.
- Castellanos, S. (2010). *Desempleo y determinación de salarios en la industria manufacturera de México: Un análisis mediante paneles dinámicos*. Econ. mex. Nueva época vol.19 no.1
- Coates, K. (2008). *Programa de Investigación Conjunta. Estimación y uso de variables no observables en la región*. México: CEMLA.
- Deossa, P. (2009). *Entrenamiento de Redes Neuronales con Filtros de Kalman*. Medellín: Universidad Nacional de Colombia.
- Diebold, F. (2004). *Elements of Forecasting*. Thomson South-Western.

- EuroStats. (2016). *Overview of GDP flash estimation methods*. Luxembourg: Publications Office of the European Union.
- EvIEWS. (2016). *User's Guide II*. Irvine CA. : IHS Global Inc.
- Ferrari, C. (2008). *Tiempos de Incertidumbre; Causas y consecuencias de la crisis mundial*. Rev.econ.inst, pp.55-78.
- FMI. (2017). *Quarterly National Accounts Manual*. FMI.
- Genschel, U. (2010). (2010). "A Comparison of Maximum Likelihood and Median Rank Regression for Weibull Estimation. Department of Statistics, State University.
- González, M. P. (2009). *Análisis de series temporales: Modelos ARIMA*. Sarriko-on.
- Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Heath, J. (19 de abril de 2012). *La Importancia de la Tasa de Desempleo*. México. obtenido de <https://jonathanheath.net/la-importancia-de-la-tasa-de-desempleo/>
- Jalles, J. T. (2009). *Structural Time Series Models and the Kalman Filter: a concise review*. UK: University of Cambridge.
- Kalman, R. (1960). *A New Approach to Linear Filtering and Prediction*. Journal of Basic Engineering, 35-45.
- Kalman, R. (1963). *New Methods in Wiener Filtering Theory*. Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability, 270-388.
- Kishor, N. K. (2012). *VAR estimation and forecasting when data are subject to revision*. Journal of Business & Economic Statistics, 181–190.
- Koopman, S. S. (1999). *Statistical algorithms for models in state space using SsfPack 2.2*. Econometrics Journal, 107-160.
- Loría, E. (2007). *Econometría con aplicaciones*. México: Pearson Educación.
- Loría, E. (2010). *Recesión y desempleo en México, un análisis estructural 1985.1-2009.2*. . IV Coloquio Internacional de Investigación en Ciencias Económico-Administrativas.

- M. Girad, A. R. (2009). *International Seminar on Timeliness, Methodology and Comparability of Economics Trends. Session 3. The availability, timeliness and quality of rapid estimates*. Ottawa.
- Maddala, G. (1977). *Econometrics*. Nueva York: McGraw-Hill.
- Mahía, R. A. (2003). *Modelos ARIMA*. Técnicas de Previsión de variables financieras.
- Meller, P. (1987). *Aplicación del Filtro de Kalman a la Estimación de Elasticidades y Variables en el Mercado de Trabajo Chileno*. Revista de Análisis Económico .
- Milas, C. F. (2008). *Out-of-sample forecasting of unemployment rates with pooled STVECM forecasts*. International Journal of Forecasting, 101–121.
- Munuera, M. (2018). *Filtro de Kalman y sus aplicaciones*. Barcelona: Universitat de Barcelona.
- Novales, A. (2014). *Modelos vectoriales autoregresivos (VAR)*. Madrid: Universidad Complutense.
- Novales, A. (1993). *Econometría*. Madrid: McGraw-Hill.
- Paganini, A. V. (2003). *Estimaciones mensuales del PBI mediante el filtro de Kalman: Evidencia de Uruguay*. Universidad de la República.
- Pierse, H. A. (1984). *Estimating Missing Observations*. Journal of the American Statistical Association 79, 125-131.
- Proietti, T. (451–476). *Forecasting the US unemployment rate*. Computational Statistics & Data Analysis, 2003.
- Raga, M. C. (2018). *Filtro de Kalman y sus aplicaciones* . Barcelona: Universitat de Barcelona.
- Serpini, G. S. (2013) *Kalmanfilter and SVR combinations in forecasting US unemployment*. Artificial Intelligence Applications and Innovations IFIP Advances in Information and Communication Technology, 506–515.
- Simionescu, M. (2015). *Kalman Filter or VAR Models*. Romanian Academy.
- Solimano, A. (1986). *Salarios reales y empleo bajo distintos regimenes macroeconomicos. Una aplicación para Chile y Brasil*. Cuadernos de Economía, Año 23, N°70, 343-371.

Spanos, A. (2002). *Teoría de la Probabilidad e Inferencia Estadística: Modelización Econométrica con Datos Observacionales*. Ciudad de México: Facultad de Ciencias, UNAM.