

Análisis de la Temperatura Mensual en Costa Rica entre 1990 y 2015 mediante regresiones en series temporales

Andrés Esteban Arguedas Leiva

Resumen

El estudio de los fenómenos climatológicos, sobre todo en los países tropicales, tiene gran importancia ya que, si es posible hacer buenas predicciones con respecto al clima a mediano plazo, se puede prever para posibles sequías, inundaciones, tormentas y huracanes. Por todo lo anterior, y sumado al papel que juega el fenómeno de *El Niño* y *La Niña* en el clima, se torna de suma importancia implementar modelos que logren hacer buenas predicciones del clima a corto o mediano plazo. Por lo anterior, este artículo presenta un análisis, mediante series de tiempo y modelos de regresión, del comportamiento de la temperatura promedio en el país, desde 1990 al 2015, con el objetivo de lograr predecir la temperatura promedio mensual con un horizonte de 3 años. Los modelos de regresión usados fueron uno con tendencia y otro con tendencia y estacionalidad, donde el modelo que tuvo los mejores resultados, tanto para el entrenamiento como para el pronóstico, es el modelo de regresión con tendencia y estacionalidad. Por último, aunque el modelo escogido hizo predicciones aceptables, es probable que este no sea el mejor modelo para este caso. Adicionalmente, puede ser útil tomar en cuenta otras covariables como la precipitación, la cantidad de CO_2 o los fenómenos de *El Niño* o *La Niña* para poder brindar mejores pronósticos.

Palabras claves: Series de tiempo, modelos de regresión, estacionalidad, temperatura, Costa Rica

Introducción

Conocer el comportamiento del clima tiene aplicaciones en muchas áreas, desde pronósticos sobre sequías e inundaciones hasta predicciones de temperatura para efectos agrícolas. Las mediciones referentes al clima tienen registros históricos bastante grandes, con mediciones que datan desde el siglo XIX. En años recientes, las mediciones sobre la temperatura y lluvias están disponibles, hasta de forma diaria, a lo largo de todo el mundo. Con estos datos, es posible crear modelos espaciotemporales para hacer predicciones del clima, no solamente a lo largo del tiempo, sino también viendo las relaciones con lugares cercanos. Este tipo de modelos son usados para hacer pronósticos climáticos a largo plazo, mediante los cuales es posible medir, por ejemplo, el posible efecto del cambio climático en distintas partes del mundo.

En el caso de Costa Rica, al ser un país tropical, se torna importante predecir ciertos eventos climáticos como, por ejemplo, *El Niño* y *La Niña*, ya que estos afectan la temperatura de los mares y, a su vez, provocan distorsiones en la temperatura, lo que afecta cultivos, pesca, producción eléctrica, entre otros. Adicionalmente, como presentó el Instituto Meteorológico Nacional (IMN) en el 2011, la cantidad de huracanes y temporales que se generan en la vertiente Caribe se disminuyen durante está el fenómeno de *El Niño*, mientras que la cantidad de huracanes en el Caribe aumenta cuando ocurre el fenómeno de *La Niña*.

Con respecto a la disponibilidad de los datos de la temperatura promedio mensual en el país, aunque estos son recopilados y procesados por el IMN, no están disponibles, en una fecha cercana, de forma abierta. Por lo anterior, la utilización de datos de otras fuentes, aunque del mismo fenómeno, se torna necesario. Al tener las mediciones de temperatura de forma mensual, es posible medir algunos patrones estacionales, o estacionalidades, que puedan existir en la temperatura. De igual forma, se pueden hacer análisis para cada año por separado, para determinar si hay años más calientes que otros, o si hay años que son diferentes a los demás. Se han realizado distintas investigaciones con respecto a pronósticos del tiempo en América Latina, como, por ejemplo, en Colombia por Mercedes Andrade (2013), donde se aplicaron modelos mixtos para hacer pronósticos climatológicos en una cierta región en Colombia. Aun así, en el caso de Costa Rica, este tipo de

análisis todavía no se han hecho de forma abierta, por lo que esta investigación trae nueva información con respecto a esta temática.

Por lo tanto, el objetivo general de este artículo es generar predicciones sobre la temperatura promedio en Costa Rica, de manera mensual y un horizonte a medio plazo, con el fin de brindar información importante para la toma de decisiones con respecto a fenómenos relacionados al clima. Adicionalmente, los objetivos específicos son describir la serie de forma descriptiva, ajustar un modelo de regresión con tendencia y con estacionalidad y, por último, hacer una predicción, con base en estos modelos, para un horizonte de tres años.

Metodología

Los datos usados en este artículo provienen del Climate Change Knowledge Portal, parte del Grupo del Banco Mundial, con base en información recolectada por el Climatic Research Unit (CRU) de la Universidad de East Anglia (UEA) en el Reino Unido. Esta información está disponible de forma gratuita en la página del CRU, tanto en forma mensual y anual, desde 1901 y hasta el 2015, en forma de una grilla de $0,5^\circ \times 0,5^\circ$ para los cuales se tiene la temperatura promedio de cada cuadro que compone la grilla. Con base en esta grilla se calcula el promedio de todas los cuadros que componen la grilla que están sobre el país y se obtiene un valor para la temperatura promedio mensual o anual para todo el territorio nacional.

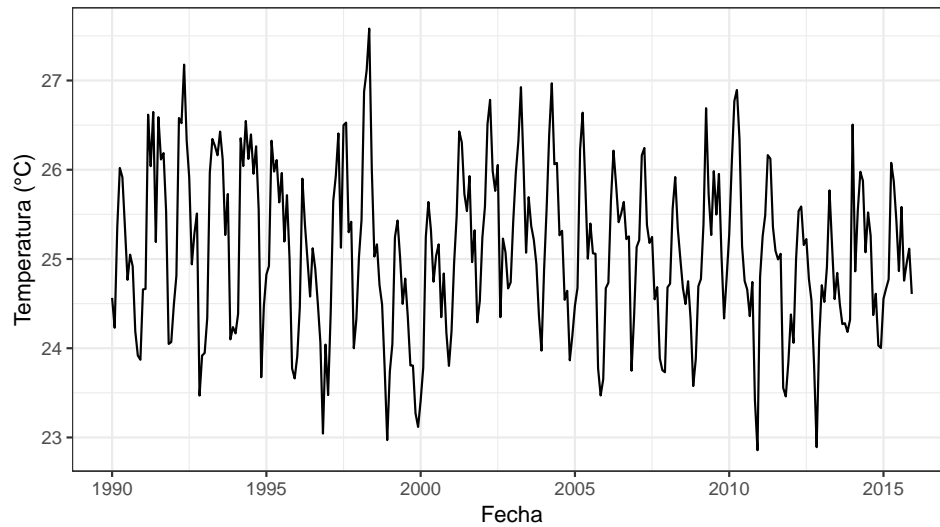
El análisis de los datos se hizo usando el lenguaje de programación R, y, de forma específica, mediante la interfaz gráfica RStudio. Adicionalmente, se usaron los paquetes *forecast*, *ggplot2*, *grid* y *gridextra*; el primero de estos para poder realizar pronósticos con los modelos de series de tiempo obtenidos y los otros tres para poder tener acceso a distintas opciones gráficas para visualizar los procedimientos y resultados.

Con respecto a los procedimientos realizados, se calcularon dos modelos de regresión, uno que tomara en cuenta la tendencia y otro que tomara en cuenta tanto la tendencia como la estacionalidad. No se realizó un modelo no lineal ya que, por el comportamiento propio de la serie, no parecía tener una tendencia no lineal que fuera fácilmente identificable, por lo que no tenía sentido hacer ese tipo de modelos. Para el cálculo de los modelos, se dividieron los datos en dos conjuntos: uno de entrenamiento y otro de prueba, este último se tomó con igual tamaño que el horizonte de pronóstico, es decir, tres años. Con base al conjunto de entrenamiento, se ajustaron ambos modelos y, posteriormente, se usaron para pronosticar los datos de prueba, comparando estos pronósticos con los datos reales para poder hacer comparaciones entre ambos modelos. Las medidas de comparación usadas son las absolutas ya que el paquete *forecast* no contiene una función para calcular las medidas de comparación relativas en las predicciones.

Resultados

Como se describió anteriormente, la serie contiene datos desde 1901 hasta el 2015, pero para efectos de este artículo solo se escogieron los datos desde Enero de 1990 hasta Diciembre del 2015, lo anterior dado que esto permite una mejor visualización de la serie y, además, por la naturaleza de los métodos a usar, se puede llegar a resultados que se vean afectados por procesos que ocurrieron hace mucho tiempo, por ejemplo, al tener una tendencia lineal, puede ser que en años recientes la serie no tenga una tendencia marcada, pero si tomamos todos los datos hay una tendencia creciente, lo cual puede generar problemas al momento de hacer pronósticos. Por lo tanto, para empezar, a continuación se presenta el gráfico de la temperatura promedio, por mes, en Costa Rica entre Enero de 1990 y Diciembre del 2015.

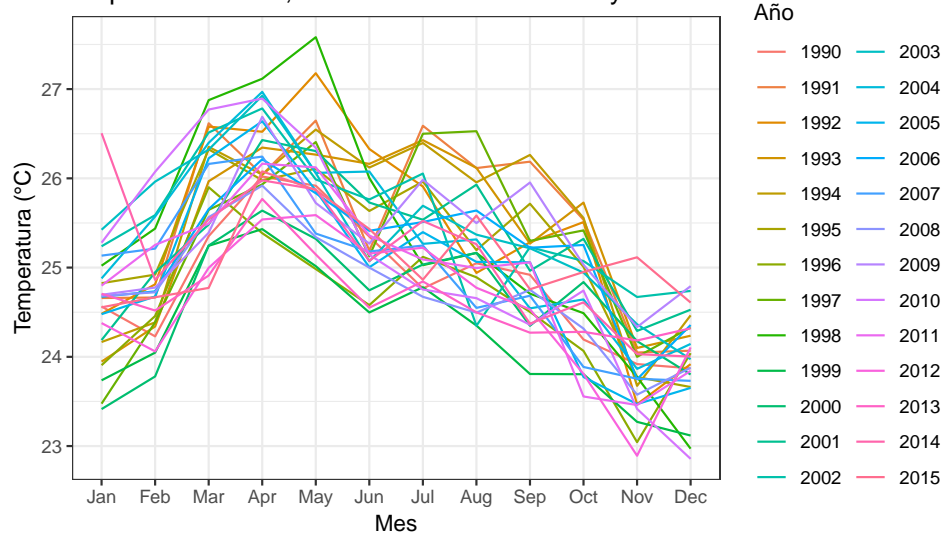
Figura 1:
Temperatura promedio por mes (en °C) entre Enero de 1990
y Diciembre del 2015, Costa Rica



Fuente: Climate Change Knowledge Portal, The World Bank Group

Como se puede notar de la Figura 1, no hay una tendencia lineal marcada, es decir, no parece crecer ni decrecer linealmente de forma notable. Aún así, parece que si hay una tendencia curva, por ejemplo, de 1990 a 1995 parece que la temperatura promedio aumenta, mientras que de 1995 al 2000 la temperatura promedio parece disminuir. Adicionalmente, los picos y valles que presenta la serie parecen tener un cierto patrón, o periodicidad, lo cual puede ser un indicio de alguna estacionalidad. Por lo anteriormente mencionado, en la Figura 2 se presenta la temperatura promedio, para cada mes, y separado según año.

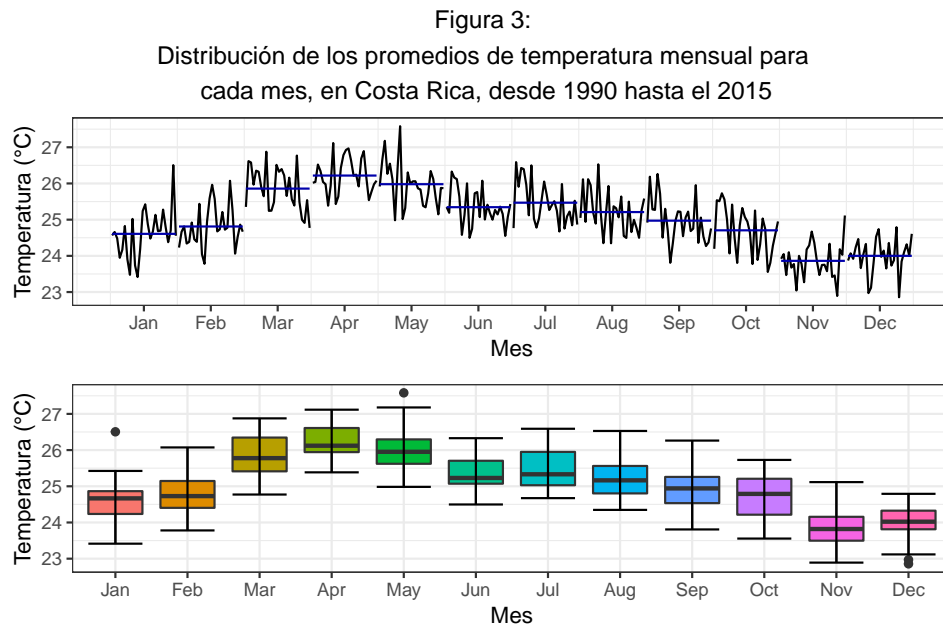
Figura 2:
Visualización de la temperatura promedio mensual según año y
para cada mes, en Costa Rica entre 1990 y 2015



Fuente: Climate Change Knowledge Portal, The World Bank Group

De la Figura 2 cabe resaltar que, aunque hay algunos casos extremos en la temperaturas promedio, como el caso que se presenta en Enero, en general para todos los años la temperatura promedio parece seguir un comportamiento bastante parecido a lo largo de los meses. Por ejemplo, se puede notar que la temperatura promedio aumenta de Enero a Marzo y se mantiene así hasta Mayo, donde empieza a disminuir y sigue

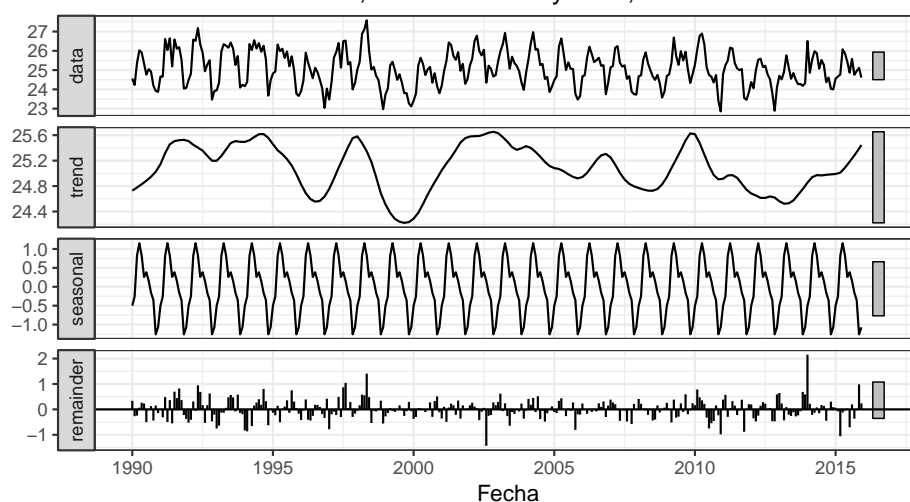
decreciendo hasta Noviembre y Diciembre, que serían los meses con temperaturas más bajas. Otra forma de visualizar esta información se presenta en la siguiente figura, que contiene, para cada mes, las temperaturas promedio para cada año y, de forma agregada, en un gráfico de cajas.



Fuente: Climate Change Knowledge Portal, The World Bank Group

Como mencionamos para la Figura 2, de la Figura 3 podemos ver la tendencia de la temperatura promedio a lo largo de los meses, lo que da indicios de que puede haber estacionalidad en la serie. Adicionalmente, cabe resaltar que la variabilidad de las temperaturas promedio para cada mes no es muy distinto, aunque si hay algunos pocos valores extremos. Por último, a nivel de cada mes, no parece haber una tendencia muy clara para todos los meses en cuanto al aumento o decrecimiento de la temperatura promedio. Por ejemplo, parece que la temperatura promedio en Enero está creciendo, mientras que la temperatura promedio en Marzo parece estar decreciendo. Por lo mencionado anteriormente, resulta útil ver la descomposición de la serie en tendencia, estacionalidad y ruido, como se presenta en la Figura 4.

Figura 4:
Descomposición de la serie de temperatura promedio mensual desde 1990 hasta el 2015 en datos, tendencia, estacionalidad y error, Costa Rica



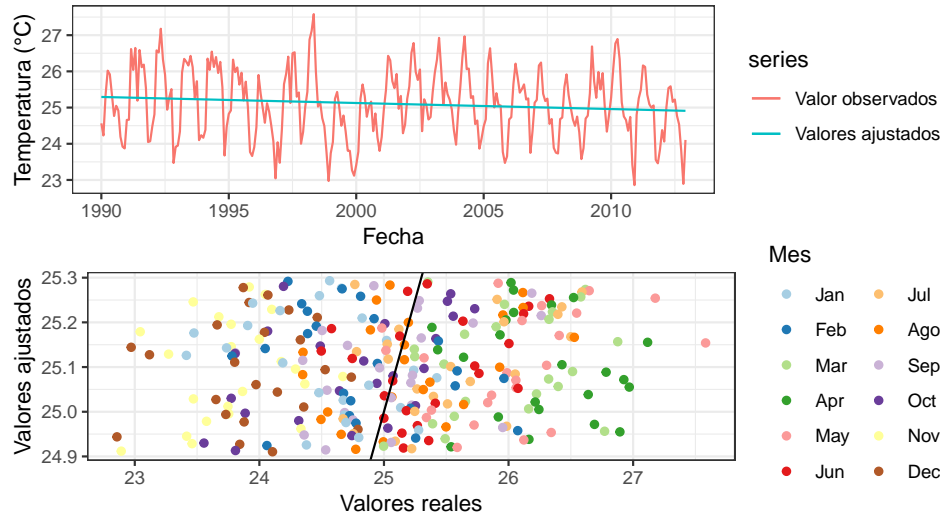
Fuente: Climate Change Knowledge Portal, The World Bank Group

De nuevo, con base en la Figura 4, parece haber una tendencia estacional en la temperatura promedio según el mes. Adicionalmente, como también se había mencionado anteriormente, la tendencia de la serie no parece ser completamente lineal, aunque si parece haber una cierta tendencia. Ahora, antes de hacer el análisis de los modelos de regresión usados, es importante clarificar que, para efectos de la medición del rendimiento de los métodos a usar, la serie se partió en dos partes, una de entrenamiento, que va de Enero de 1990 a Diciembre del 2012, y una de prueba, que empieza en Enero del 2013 y termina en Diciembre del 2015. Por lo tanto, la serie de entrenamiento contiene información de 13 años y la de prueba de 3 años, lo anterior con el objetivo de poder medir el rendimiento en predicción de los modelos a generar, dado que el horizonte a usar es de 3 años para las predicciones finales. Habiendo hecho la partición de la serie, se procede a ajustar un modelo de regresión cuya única variable sea la tendencia. El Cuadro 1 presenta los coeficientes del modelo, sus errores estándar y las probabilidades asociadas a estos, mientras que la Figura 5 presenta la visualización del modelo sobre la serie y los valores ajustados mediante el modelo contra los observados en la serie.

Cuadro 1:
Coeficientes, errores estándar y probabilidades asociadas para el modelo de regresión con tendencia

Variable	Coeficiente	Error Estándar	Pr(> t)
Intercepto	25.295	0.11	<0.001
Tendencia	-0.001	0.001	0.049

Figura 5:
Valores observados y ajustados mediante una regresión con tendencia de la temperatura promedio mensual en Costa Rica, desde 1990 hasta el 2012



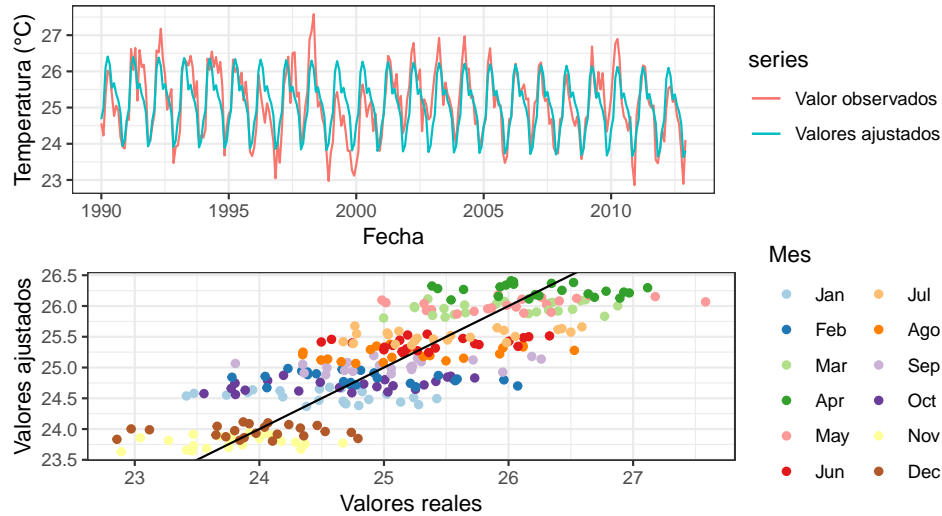
Fuente: Climate Change Knowledge Portal, The World Bank Group

Aunque el modelo de regresión con tendencia logra ajustar, como se presenta en la Figura 5, este ajuste no parece ser especialmente bueno, ya que no logra captar los picos y valles presentes en la serie. Además, al analizar el gráfico de valores ajustados contra observados, se puede notar que hay una dispersión casi aleatoria y no hay acumulación de puntos cerca de la línea, lo que significa que el ajuste no es realmente adecuado. Luego, con respecto a lo obtenido en el Cuadro 1, se pueden resaltar dos aspectos. El primero es que la temperatura promedio mensual en Costa Rica, entre 1990 y 2012 es de 25,295, y el segundo es que la variable de la tendencia resulta significativo, aunque no por mucho, y es negativo, lo que indica que, en promedio, la temperatura promedio mensual en Costa Rica está decreciendo. Ahora, como se comentó con base en algunas figuras presentadas anteriormente, hay indicios de que existe una cierta estacionalidad. Por lo anterior, tiene sentido hacer un modelo de regresión que tome en cuenta tanto la tendencia como la estacionalidad. Los resultados de calcular este modelo se presentan en el Cuadro 2 y la Figura 6, donde se muestran los coeficientes junto con algunos estadísticos relacionados y el ajuste del modelo a la serie, respectivamente.

Cuadro 2:
Coeficientes, errores estándar y probabilidades asociadas para el modelo de regresión con tendencia y estacionalidad

Variable	Coeficiente	Error Estándar	Pr(> t)
Intercepto	24.68	0.13	<0.001
Tendencia	-0.001	0.001	0.006
Estacionalidad			
Febrero	0.306	0.166	0.066
Marzo	1.439	0.166	<0.001
Abril	1.735	0.166	<0.001
Mayo	1.506	0.166	<0.001
Junio	0.869	0.166	<0.001
Julio	1.002	0.166	<0.001
Agosto	0.707	0.166	<0.001
Setiembre	0.524	0.166	0.002
Octubre	0.205	0.166	0.219
Noviembre	-0.725	0.166	<0.001
Diciembre	-0.551	0.166	0.001

Figura 6:
Valores observados y ajustados mediante una regresión con tendencia y estacionalidad
de la temperatura promedio mensual en Costa Rica, desde 1990 hasta el 2012



Fuente: Climate Change Knowledge Portal, The World Bank Group

Comparando los resultados de la Figura 6 con los de la Figura 5, podemos notar que el ajuste del modelo usando tanto la tendencia como la estacionalidad parece ser bastante mejor al del modelo con solo la tendencia. Además de lo anterior, al comparar el gráfico de valores ajustados contra observados, se puede observar que para el modelo con tendencia y estacionalidad, los valores se aglomeran en mayor cantidad sobre la línea de identidad y están separados claramente según el mes, lo que indica un buen ajuste del modelo. Luego, con respecto al Cuadro 2, cabe resaltar el coeficiente de la tendencia es significativo y negativo, al igual que en el modelo con solo la tendencia, mientras que nueve de los once coeficientes de estacionalidad son significativos, por lo que parece que si fue útil haberlo tomado en cuenta en el modelo. Adicionalmente, es importante recordar que el intercepto cambió con respecto al intercepto del modelo con solo la tendencia, ya que el intercepto en el modelo con tendencia y estacionalidad se refiere a la temperatura promedio en enero, y los coeficientes de estacionalidad para cada mes se refieren a cuanto aumenta, o disminuye, la temperatura para un cierto mes con respecto a enero. A continuación, para poder comparar ambos modelos en términos de predicción, la Figura 7 presenta los pronósticos para tres años obtenidos de ambos modelos, mientras que el Cuadro 3 presenta algunas medidas de rendimiento absolutas para comparar ambos modelos.

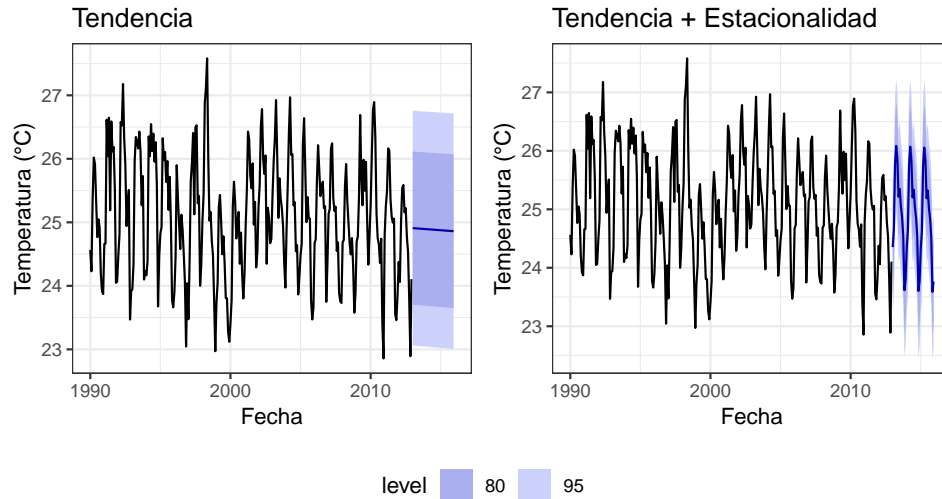
Cuadro 3:

Medidas de rendimiento para los conjuntos de entrenamiento y prueba para los dos modelos de regresión

Modelo	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	U de Theil
Tendencia								
Entrenamiento	0.000	0.929	0.761	-0.138	3.040	1.300	0.700	
Prueba	0.084	0.625	0.498	0.277	1.979	0.852	0.403	0.935
T + Estacionalidad								
Entrenamiento	0	0.549	0.447	-0.048	1.778	0.763	0.614	
Prueba	0.051	0.619	0.446	0.190	1.782	0.762	0.392	0.930

Fuente: Climate Change Knowledge Portal, The World Bank Group

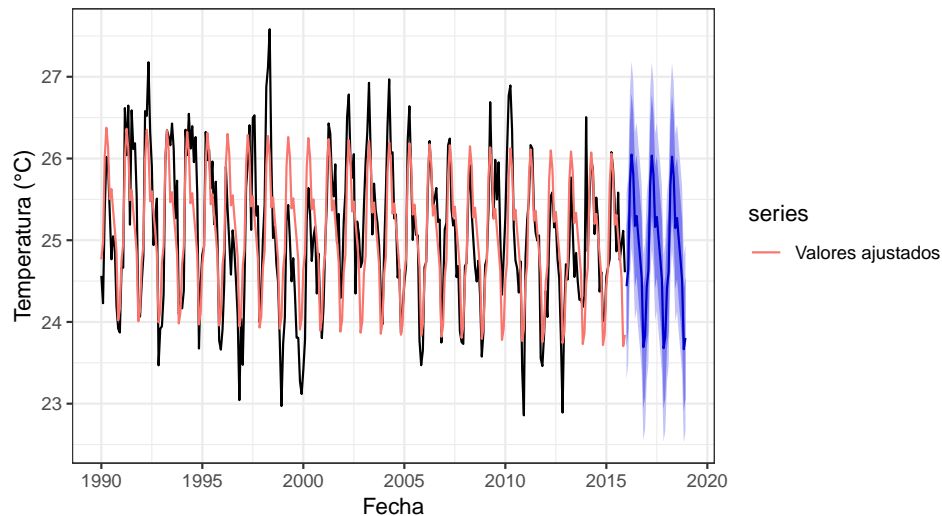
Figura 7:
Valores predichos para ambos tipos de regresión y sus intervalos de confianza,
para la temperatura promedio en Costa Rica del 2013 al 2015



Fuente: Climate Change Knowledge Portal, The World Bank Group

Los resultados de las predicciones obtenidos en la Figura 7 son los esperados con base en lo analizado de ambos modelos para el conjunto de entrenamiento. Por esto, el modelo con solo tendencia hace pronosticos en una línea recta con un intervalo de confianza bastante ancho, mientras que el modelo con tendencia y estacionalidad parece seguir el comportamiento estacional de la serie, además de que tiene intervalos de confianza más estrechos. Ahora, con respecto al Cuadro 3, el modelo con tendencia y estacionalidad tiene todos los indicadores menores que el modelo con sólo la tendencia, tanto para los datos de entrenamiento como los de prueba. Por todo lo mencionado anteriormente, parece que el modelo que mejor ajusta de los presentados es el modelo regresión con tendencia y estacionalidad. Por lo tanto, en la Figura 8 se presentan las predicciones para la temperatura promedio mensual desde Enero del 2016 hasta Diciembre del 2018, con base en un modelo de regresión con tendencia y estacionalidad, es decir, predicciones a un horizonte de 36 meses, o 3 años.

Figura 8:
Valores observados y pronosticados mediante una regresión con tendencia y estacionalidad
de la temperatura promedio mensual en Costa Rica, desde 1990 hasta el 2018



Fuente: Climate Change Knowledge Portal, The World Bank Group

Conclusiones

Como principal conclusión, se logró obtener un modelo que logra predecir, bajo un horizonte de tres años, la temperatura mensual promedio tomando en cuenta tanto la tendencia como la estacionalidad. Este modelo se ajustó a un conjunto de entrenamiento y se predijo sobre un conjunto de prueba, obteniendo así medidas de rendimiento que fueron aceptables. Además, es importante recordar que el estudio de fenómenos climatológicos en Costa Rica se torna importante ya que afecta muchas áreas productivas y de la vida diaria de los habitantes, por lo que el análisis mediante series de tiempo de estos fenómenos brinda información importante para la toma de decisiones.

Consecuentemente, relacionado a lo obtenido en los resultados, es importante resaltar que, como era de esperar, el modelo de regresión que usaba solamente la tendencia no dio resultados muy buenos, mientras que los resultados del modelo con tendencia y estacionalidad fueron mejores. Aún así, el ajuste de este modelo no es perfecto y hay ciertos picos y valles en la serie que el modelo no logra identificar de buena manera, por lo que probablemente al replicar el análisis con otros modelos, el ajuste de la regresión sea peor. Por lo anterior, no parece que el modelo de regresión sea la mejor opción para modelar este fenómeno y, para futuras investigaciones, sería interesante ver el efecto que tiene “El Niño” y “La Niña” en las estimaciones de estos modelos. Adicionalmente, como se menciona tanto durante el análisis descriptivo de la serie como en el análisis de los modelos, la serie parece tener una estacionalidad bastante marcada, por lo que si se quiere aplicar algún otro modelo a esta serie, lo recomendable sería que este tome modelo logre tomar en cuenta la estacionalidad de la serie.

Por último, con respecto al modelo y al pronóstico obtenido, es importante notar que las predicciones que brinda el modelo son para temperaturas promedio mensuales esperadas, por lo que, si una temperatura para un cierto mes está fuera de los intervalos, esto puede ser un indicio de que está ocurriendo un fenómeno del Niño, si la estimación es superior a los límites, o la Niña si es menor a los límites. Adicionalmente, también cabe destacar que los intervalos de confianza para la regresión resultan ser bastante anchos, pero no aumentan mucho a lo largo del tiempo, por lo que, para pronósticos en corto a mediano plazo, estos no deberían estar muy lejos de los valores que se van a observar. Por otro lado, por la forma misma en la cual se definió el modelo, podemos notar que las predicciones siguen una tendencia estacional bastante marcada, pero casi constante a lo largo de los años, lo que nos indica que esperamos que, a corto o mediano plazo, la tendencia general de la temperatura mensual no cambie de forma drástica.

Bibliografía

- Andrade, M. (2013). *Monthly Average Temperature Modeling in an Intertropical Region*. American Meteorologist Society 28 (5), 1099-1115.
- Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. Extraído de <https://CRAN.R-project.org/package=gridExtra>
- Instituto Meteorológico Nacional. *El ENOS y su efecto en Costa Rica*.
- Hyndman, R. J. & Khandakar, Y. (2008). *Automatic time series forecasting: the forecast package for R*. Journal of Statistical Software 26 (3), 1-22. Extraído de <http://www.jstatsoft.org/article/view/v027i03>
- R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Extraído de <https://www.R-project.org/>
- RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, Massachusetts. Extraído de <http://www.rstudio.com/>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.