



UNIVERSIDAD MICHOCANA DE SAN NICOLAS DE HIDALGO

FACULTAD DE INGENIERÍA ELÉCTRICA

TESIS

**PREDICCIÓN DE SERIES DE TIEMPO CON EL MODELO ARMA A
TRAVÉS DE ALGORITMOS GÉNETICOS.**

**QUE PARA OBTENER EL TÍTULO DE:
INGENIERO EN COMPUTACIÓN**

**PRESENTA:
VICTOR MANUEL GARCÍA VILCHEZ**

**ASESOR:
DR. JUAN JOSÉ FLORES ROMERO**

MORELIA, MICH.A JUNIO DEL 2011



**FACULTAD DE
INGENIERÍA ELÉCTRICA**

Agradecimientos

*Agradezco a mi asesor por el apoyo brindado
para esta tesis así como sus consejos*

Para llevar a cabo la misma.

*Agradezco a mis padres por el apoyo
y la confianza que tuvieron
en mi para salir adelante.*

*Agradezco a mis compañeros,
amigos y profesores de la Facultad
por brindarme su amistad
y sus enseñanzas mismas que me
sirvieron durante mi carrera.*

Dedicatoria

*A mis padres y amigos que siempre me dieron
esperanzas para terminar la carrera y
me apoyaron para no rendirme a
medio camino.*

Resumen

El clima es un factor importante en nuestras vidas. Si se pudiera saber cuánto frío o calor hará en días futuros como mañana o pasado mañana o incluso lo de todo un mes se podría usar para evitar enfermedades de gripa y así saber si salir abrigado o no. Ahora los pronósticos son más exactos. Para hacer una predicción se debe contar con datos, observaciones o valores medidos en un tiempo determinado, los cuales forman parte de la serie de tiempo. Aún no hay modelos precisos que digan con exactitud el valor exacto de qué pasará mañana. Por eso en esta tesis se manejará el modelo ARMA (Autorregresivo de Media Móvil) y Algoritmos Genéticos, los cuales deberán manipular los datos y darnos valores aproximados a la serie de tiempo original.

Para manejar los Algoritmos Genéticos, los cuales fueron inspirados en el origen de las especies de Charles Darwin [Darwin 1859], se usará Mathematica y Evolvica [Jacob 1987].

Evolvica fue desarrollada en Mathematica, y simplemente es una implementación de Algoritmos Genéticos, los cuales manejan muchos conceptos de la biología, por ejemplo cromosoma, combinación, mutación, etc. Los Algoritmos Genéticos siempre hacen una búsqueda de optimización; lo importante es saber cómo interpretar los datos en la computadora para así lograr tener una traducción de los mismos. Con esto podemos ver que tan bueno es un individuo, lo cual se conoce como función de aptitud. Esta función nos dice que tan bueno es un individuo el cual puede ser un cromosoma de números o letras. Para esta tesis nuestros cromosomas serán números reales y tendrán dos tipos de rango, (los cuales serán especificados en el capítulo 4) y serán escogidos aleatoriamente; después se descompondrán para formar parte de la nueva serie de tiempo que se creará con el Modelo Autorregresivo de Media Móvil. Después de crearse la serie de tiempo se comparará con la original y nos dirá cuánto error tiene y además se especificará un error mínimo para saber si es apto o no. Solo los cromosomas aptos serán graficados y solo se mostrará el peor y el mejor individuo. Los individuos que se crean dependen del número de marcos que se manejarán para decir que tan buenos son los Algoritmos Genéticos con respecto al modelo ARMA y cuales se recomiendan más para quedarse a seguir buscando. En esa parte se tomará una selección del individuo más apto, entre un grupo de

cromosomas. Al individuo más apto se le practicarán operaciones de combinación y mutación. También se especificará en los experimentos cuantos individuos deberán participar y serán 3 experimentos por cada tamaño del marco, el cual representa el número de valores para construir el modelo ARMA.

Contenido

| | |
|-----------------------------------------------------------------------------------------|-----------|
| Agradecimientos..... | ii |
| Dedicatoria..... | iii |
| Resumen..... | iv |
| Contenido..... | vi |
| Lista de figuras..... | viii |
| Lista de tablas..... | x |
| Lista de símbolos y abreviaciones..... | xi |
| | |
| Capítulo 1. Introducción..... | 1 |
| 1.1. Antecedentes..... | 2 |
| 1.2. Objetivos de la tesis..... | 3 |
| 1.3. Justificación..... | 3 |
| 1.4. Metodología..... | 4 |
| 1.5. Contenido de la tesis..... | 4 |
| | |
| Capítulo 2. Conceptos básicos de la series de tiempo..... | 5 |
| 2.1 Modelos Box-Jenkins..... | 5 |
| 2.2 Modelo autorregresivo de media móvil..... | 7 |
| 2.3 Algoritmos genéticos..... | 10 |
| | |
| Capítulo 3. Desarrollo de la programación de series de tiempo con Mathematica... | 12 |
| | |
| Capítulo 4. Experimentos con algoritmos genéticos para series de tiempo..... | 23 |
| 4.1 Experimentos con un marco de 12..... | 25 |
| 4.2 Experimentos con un marco de 18..... | 30 |
| 4.3 Experimentos con un marco de 24..... | 35 |

| | |
|--------------------------------------|-----------|
| Capítulo 5. Conclusiones..... | 40 |
| Referencias..... | 42 |
| ANEXO..... | 43 |
| Valores del capítulo 3..... | 43 |

Lista de Figuras

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 Figura 1.1 Ejemplos de diversas series de tiempo (a) accidentes de tránsito de Morelia y (b) Promedio Industrial Dow Jones..... | 1 |
| 2.1 Gráfica de los datos de la tabla 2.1..... | 6 |
| 2.2 Serie de Tiempo con un Marco (k); el cual indica el número de valores de la gráfica que serán tomados en consideración..... | 8 |
| 2.3 Serie de Tiempo con seis valores calculados a través del modelo ARMA..... | 10 |
| 2.4 Ejemplos de diversos tipos de Recombinación (a) combina la mitad de los padres para los nuevos hijos, (b) combina de diferentes lados para los padres y obtener hijos muy combinados..... | 11 |
| 2.5 Selección (a) lugar donde se situará la población (b) se ve la población en diversas partes así como diversos individuos caracterizados por su color, (c) se selecciona a los individuos más en aptos..... | 12 |
| 3.1 Gráfica de temperatura promedio mensual de la ciudad de Morelia, de 1° Enero del 2010 hasta el 15 de Octubre del 2010..... | 14 |
| 3.2 Demostración del uso de corchetes dobles sobre un cromosoma (a) representa el primer elemento del chr1 mientras que (b) muestra el segundo elemento del chr1..... | 16 |
| 3.3 Comparación de Gráficas de temperatura promedio mensual de la ciudad de Morelia, original y con el modelo ARMA generado por la función f..... | 20 |
| 4.1 Serie de Tiempo de temperatura promedio mensual en la ciudad de Morelia..... | 24 |
| 4.2 Serie de Tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 12 valores del marco inicialmente..... | 25 |
| 4.3 Resultados obtenidos con un tamaño de marco de 12 valores..... | 26 |

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.4 Representación de la serie de tiempo original y el modelo ARMA de la Ecuación (4.1)..... | 28 |
| 4.5 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 1.0655 °C..... | 29 |
| 4.6 Serie de Tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 18 valores del marco inicialmente..... | 30 |
| 4.7 Resultados obtenidos con un tamaño de marco de 18 valores..... | 31 |
| 4.8 Representación de la serie de tiempo original y el modelo ARMA de la Ecuación (4.4)..... | 33 |
| 4.9 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 0.2897 °C..... | 34 |
| 4.10 Serie de Tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 24 valores del marco inicialmente..... | 35 |
| 4.11 Resultados obtenidos con un tamaño de marco de 36 valores..... | 36 |
| 4.12 Representación de la serie de tiempo original y el modelo ARMA de la Ecuación (4.4)..... | 38 |
| 4.13 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 0.2325 °C..... | 39 |

Lista de Tablas

| | |
|-----------------------------------------------------------------------------------|----|
| 2.1 Datos de desempleados..... | 5 |
| 4.1 Datos obtenidos de los experimentos con un tamaño de marco de 12 valores..... | 27 |
| 4.2 Datos obtenidos de los experimentos con un tamaño de marco de 18 valores..... | 32 |
| 4.3 Datos obtenidos de los experimentos con un tamaño de marco de 24 valores..... | 37 |

Lista de Símbolos y Abreviaturas

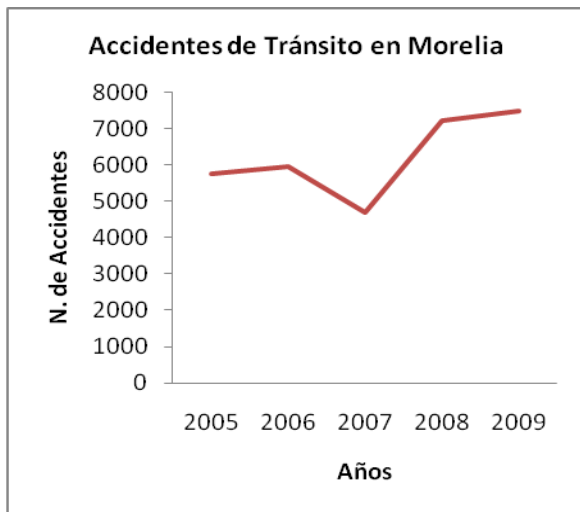
| | |
|---------|---------------------------------------------------------|
| AG | Algoritmos Genéticos |
| ARMA | Autorregresivo de Media Móvil |
| ARIMA | Autorregresivo Integrado de Media Móvil |
| AR | Autorregresivo |
| MA | Media Móvil |
| K | Número de Marco |
| X_t | Valores de la Serie de Tiempo |
| ξ_t | Errores de la serie de tiempo |
| a_i | Valores de a del modelo ARMA |
| b_i | Valores de b del modelo ARMA |
| E | Valor de c del modelo ARMA |
| p | Marco de números para la serie de tiempo |
| q | Marco de números para los errores de la serie de tiempo |

Capítulo 1

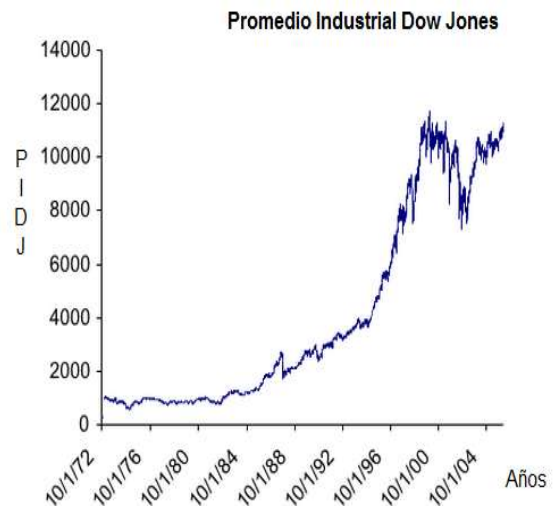
Introducción

La predicción de series de tiempo [STASTIK 2010] se realiza con datos históricos, estos pueden ser datos climatológicos, de las acciones de la bolsa de valores, de recursos de una empresa, de accidentes ocurridos en carreteras, de inflación, desempleo, etc. Las series de tiempo nos describen el comportamiento de distintos fenómenos.

Es una gran importancia pronosticar el comportamiento de una variable para hacer una planeación [Brockwell 1987]. En la Figura 1.1 se observan distintas series de tiempo; en (a) muestra el número de accidentes ocurridos en la ciudad de Morelia desde el año 2005 hasta el 2009; en (b) muestra la agrupación de las 30 compañías estadounidenses más importantes mejor conocida como Promedio Industrial Dow Jones del 10 de enero de 1972 hasta el 10 de enero del 2004.



(a)



(b)

Figura 1.1 Ejemplos de diversas series de tiempo (a) accidentes de tránsito de Morelia y (b) Promedio Industrial Dow Jones.

Los métodos de predicción nunca serán exactos pero nos ayudan imaginar cómo podría ser el comportamiento de la serie de tiempo que estemos manejando, para ello se utilizará en esta tesis el modelo ARMA (Modelo Autorregresivo de Media Móvil) [Brockwell 1987], el cual estará implementado con Algoritmos Genéticos (AG) [Wellin 2004]. Los AG son parte de una rama de la Inteligencia Artificial y fueron inspirados en la evolución biológica, los cuales toman datos al azar creando un individuo con ciertas características. Para la realización de los algoritmos genéticos se utilizará un software llamado Mathematica [Ruskeepää 1987] y Evolvica [Jacob 2001].

Mathematica [Wellin 2005] es un software que realiza cálculos matemáticos y de programación; su ventaja más grande es el uso de manipulaciones simbólicas.

Evolvica [Jacob 2001] fue desarrollado en Mathematica por Cristian Jacob el cual nos permitirá manejar cromosomas y con ellos poder crear individuos, los cuales serán manipulados para obtener resultados concretos en nuestro caso, series de tiempo.

1.1 Antecedentes

Una red neuronal artificial es un modelo computacional que funciona igual que las redes neuronales biológicas; conocidas para modelar relaciones complejas entre la entrada y salida.

Entre esas conexiones están las capas ocultas también llamadas neuronas, la búsqueda de los parámetros se vuelve enorme. Las neuronas también son conocidas como capas las cuales se conectan entre si, además de que los algoritmos de entrenamiento que se implementan a las redes neuronales artificiales dan un resultado similar a algoritmos genéticos.

Se ha utilizado a la par los dos métodos, sin embargo no es muy común el uso de esta combinación ya que cuesta trabajo unirlos para obtener una optimización de recursos.

Los modelos que manejan predicciones pasadas y errores, son los modelos de Box-Jenkins los cuales utilizan los movimientos Autorregresivos como ARIMA o ARMA. Estos modelos en la actualidad se utilizan para series Económicas, Físicas, Demográficas, Marketing, Telecomunicaciones o de Transporte [Brockwell 1987].

El diseño de los modelos de predicción de series de tiempo tradicionalmente se ha realizado mediante métodos estadísticos. Es difícil modelar series de tiempo, pero el

modelo ARMA tiene la ventaja por la manera en la que se utiliza ya que va corrigiendo el error de la predicción.

Los Algoritmos Genéticos surgieron en 1970 propuestos por J. H. Holland, fueron inspirados en la evolución biológica y base genético-molecular; estos algoritmos hacen evolucionar una población de individuos los cuales son generados de forma aleatoria y deben ser sometidos a una selección para obtener el individuo más apto y mejor.

Aunque el concepto de Algoritmos Genéticos nació en los 70's es apenas en años recientes que se ha difundido o popularizado su uso; ya que son parte de la inteligencia artificial. La inteligencia artificial se ha convertido en una alternativa importante en la resolución de problemas de optimización de recursos.

1.2 Objetivo

El objetivo de esta tesis radica en encontrar por medio del modelo ARMA y Algoritmos Genéticos el pronóstico en la serie de tiempo del clima de la ciudad de Morelia en el estado de Michoacán.

1.3 Justificación

Se manejan los algoritmos genéticos los cuales se les han hecho a un lado por parte de los alumnos de la Facultad de Ingeniería Eléctrica, a falta de una motivación tanto de los profesores como de los estudiantes. Por tal motivo esta tesis demostrara que se puede tener una respuesta adecuada en la optimización de la predicción de series de tiempo.

Por ello se escogió la temperatura promedio mensual de la ciudad de Morelia para que sean más claros los algoritmos genéticos en la Facultad. También se desea que este trabajo motive a que futuras generaciones que estudian la carrera de Ingeniería en Computación y realicen trabajos de optimización con algoritmos genéticos.

1.4 Metodología

La metodología que se utilizará para llegar al objetivo de la tesis será un modelo ARMA, el cual será auxiliado por algoritmos genéticos. Para ello se define un cromosoma donde los valores que se generán son los coeficientes del modelo ARMA y los cuales forman una nueva serie de tiempo que será el individuo a evaluar con la función de aptitud.

El uso de algoritmos genéticos requiere de una población inicial la cual evolucionará a partir de ciertos parámetros que se deben definir como el número de generaciones, el número de padres, el número de hijos, el número de combinaciones, el número de cromosomas, el tipo de alfabeto del cromosoma y la función de aptitud.

Cuando se han definido todos los parámetros, a la población inicial se le realiza una selección de los individuos más aptos por medio de la función de aptitud. Los mejores individuos son sometidos a las operaciones de recombinación y mutación.

La recombinación requiere de dos cromosomas para tener a uno nuevo, mientras que la mutación solo necesita un cromosoma para formar uno nuevo. Ambos operadores se utilizan para tener un individuo mejor de cada generación.

Al terminar la evolución de la población inicial se tendrá el mejor individuo de todas las generaciones se le realizará una predicción para ver que tan acertado es el modelo ARMA obtenido del experimento.

1.5 Contenido de la tesis

En el capítulo 1 se da una introducción a este trabajo, se muestra la situación actual de las series de tiempo y se dan a conocer el modelo y software que se utilizarán a lo largo de la tesis.

En el capítulo 2 se presentan conceptos básicos de la teoría de series de tiempo, así como la aplicación del modelo ARMA. Se describe la formulación de la serie de tiempo del clima de la ciudad de Morelia del estado de Michoacán del país de México.

En el capítulo 3 se presentan el desarrollo de la programación que llevara así como la creación de individuos que serán parte de los futuros experimentos que se realizarán.

En el capítulo 4 se abordan los experimentos dándonos resultados y observaremos cual será el mejor de nuestros experimentos que serán hechos a través de evolución.

En el capítulo 5 se presentan conclusiones generales, aportaciones y trabajos futuros.

Capítulo 2

Conceptos básicos de series de tiempo

Una Serie de Tiempo se crea a través de una secuencia de observaciones, secuencias de datos, de valores en determinados instantes de tiempo los cuales están distribuidos cronológicamente y espaciados entre sí de manera uniforme [Brockwell 1987].

Existen muchas razones para hacer registros y análisis de datos de una serie de tiempo puesto que entre los datos nos permiten obtener una mejor comprensión del mecanismo de generación de datos, la predicción de futuros valores y así poder tener un control óptimo de dichos valores. Una propiedad característica de las series de tiempo [STASTIK 2010] es que los datos no se generan de forma independiente, su dispersión (distancia entre dos puntos o más) varía en el tiempo, a menudo se rigen por una tendencia y tienen componentes periódicos. Los procedimientos estadísticos son dependientes de los análisis de series de tiempo ya que se requiere de métodos adecuados para obtener dicha predicción de valores. Los datos de la Tabla 2.1 son el número de trabajadores desempleados en el sector de la construcción de Alemania desde Julio de 1975 hasta septiembre de 1979 [STASTIK 2010]. La Figura 2.1 muestra gráficamente el contenido de la tabla 2.1.

2.1 Modelos Box-Jenkins

El nombre de este método viene de George Box y Jenkins Gwilym los cuales propusieron la aplicación de los modelos Autorregresivos de media móvil como lo son el ARMA y ARIMA.

Los modelos Box-Jenkins utilizan un método iterativo de tres etapas de modelado:

- **Identificación y Selección del Modelo:** En este paso es necesario saber si las variables son estacionarias, determinar el tamaño de marco para los parámetros p que es la cantidad de elementos del modelo autorregresivo y q es la cantidad de elementos de la media móvil del modelo ARMA para generar una en la serie de tiempo. Si una serie de tiempo es estacionaria se aplica el modelo ARMA en caso contrario se aplica ARIMA.

Tabla 2.1 Datos de desempleados

| N | Mes | Desempleados | N | Mes | Desempleados | N | Mes | Desempleados |
|----|------------|--------------|----|------------|--------------|----|------------|--------------|
| 1 | Julio | 60572 | 18 | Diciembre | 65036 | 35 | Mayo | 33990 |
| 2 | Agosto | 52461 | 19 | Enero | 110728 | 36 | Junio | 26819 |
| 3 | Septiembre | 47357 | 20 | Febrero | 108931 | 37 | Julio | 25291 |
| 4 | Octubre | 48320 | 21 | Marzo | 71517 | 38 | Agosto | 24538 |
| 5 | Noviembre | 60219 | 22 | Abril | 54428 | 39 | Septiembre | 22685 |
| 6 | Diciembre | 84418 | 23 | Mayo | 42911 | 40 | Octubre | 23945 |
| 7 | Enero | 119916 | 24 | Junio | 37123 | 41 | Noviembre | 28245 |
| 8 | Febrero | 124350 | 25 | Julio | 33044 | 42 | Diciembre | 47017 |
| 9 | Marzo | 87309 | 26 | Agosto | 30755 | 43 | Enero | 90920 |
| 10 | Abril | 57035 | 27 | Septiembre | 28742 | 44 | Febrero | 89340 |
| 11 | Mayo | 39903 | 28 | Octubre | 21968 | 45 | Marzo | 47792 |
| 12 | Junio | 34053 | 29 | Noviembre | 41427 | 46 | Abril | 28448 |
| 13 | Julio | 29905 | 30 | Diciembre | 63685 | 47 | Mayo | 19139 |
| 14 | Agosto | 28068 | 31 | Enero | 99189 | 48 | Junio | 16728 |
| 15 | Septiembre | 26634 | 32 | Febrero | 104240 | 49 | Julio | 16523 |
| 16 | Octubre | 29259 | 33 | Marzo | 75304 | 50 | Agosto | 16622 |
| 17 | Noviembre | 38942 | 34 | Abril | 43622 | 51 | Septiembre | 15499 |

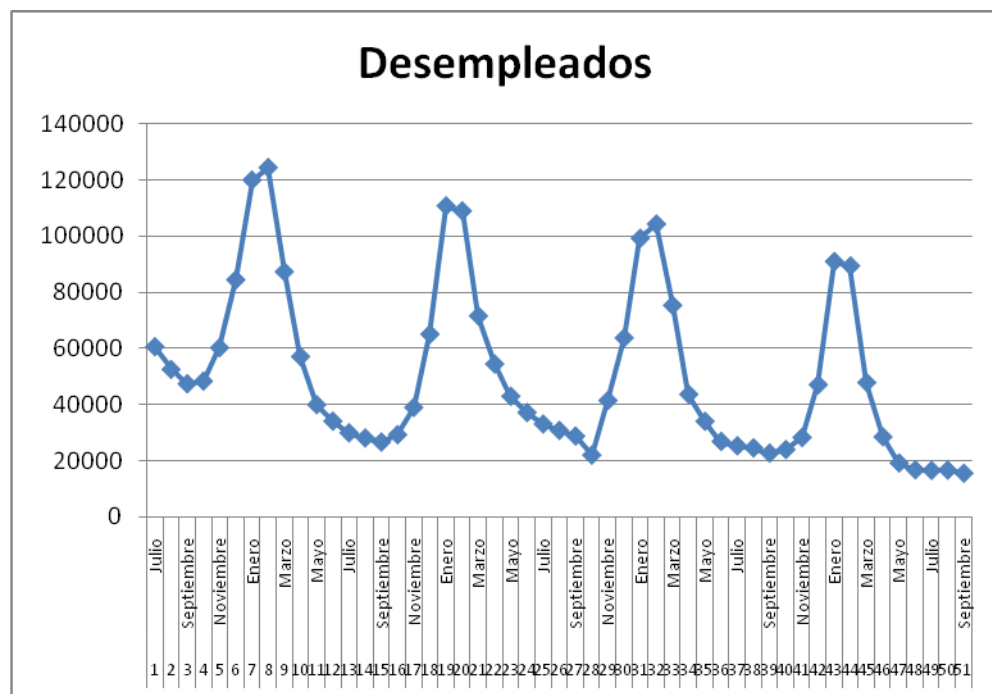


Figura 2.1 Gráfica de los datos de la tabla 2.1.

- Estimación de los parámetros: usa los algoritmos para estimar los coeficientes que seleccionen el modelo más adecuado entre los modelos ARMA y ARIMA. Los métodos más comunes buscan el óptimo en las estimaciones.
- Modelo de control: sirve para comprobar si el modelo estimado se ajusta a las especificaciones de un proceso univariante estacionario (las observaciones son ordenadas y equidistantes cronológicamente).

2.2 Modelo autorregresivo de media móvil (ARMA)

El modelo está formado por dos partes, una es Autorregresivo (AR) y la otra de Media Móvil (MA). El modelo ARMA (p, q) maneja dos variables: p es la cantidad de elementos $a_1 \dots a_p$ para el modelo autorregresivo que se maneja en la serie de tiempo; q es la cantidad de elementos $b_1 \dots b_q$ para la media móvil la cual se encarga de mejorar el modelo autorregresivo.

El Modelo Autorregresivo (AR (p)) se puede escribir como se indica en la Ecuación (2.1):

$$\hat{x}_t = E + \sum_{i=1}^p a_i x_{t-i} \quad (2.1)$$

donde $a_1 \dots a_p$ son los parámetros del modelo, E es una constante y $x_{t-1} \dots x_{t-p}$ son datos de la serie de tiempo.

La notación MA (q) se refiere a un modelo de Media Móvil de orden q como se ve en la Ecuación (2.2):

$$\hat{x}_t = \sum_{i=1}^q b_i \xi_{t-i} \quad (2.2)$$

donde $b_1 \dots b_q$ son los parámetros del modelo y $\xi_{t-1} \dots \xi_{t-q}$ son los términos de error.

La notación ARMA (p, q) se refiere a un modelo con p términos Autorregresivos y q términos de Media Móvil por lo cual combinados obtendremos la siguiente Ecuación (2.3)

$$\hat{x}_t = \sum_{i=1}^p a_i x_{t-i} + \sum_{i=1}^q b_i \xi_{t-i} + E \quad (2.3)$$

donde $a_1 \dots a_p$ y $b_1 \dots b_q$ son los parámetros del modelo, E es una constante, $\xi_{t-1} \dots \xi_{t-q}$ son términos de error y $x_{t-1} \dots x_{t-p}$ son datos de la serie de tiempo.

En la Ecuación (2.3) se observa el modelo ARMA, el cual nos ayudará a predecir el siguiente valor, con la cual estaremos armando una serie de tiempo. Para conseguirlo lo podemos usar de dos maneras una se puede considerar con p y q donde ambas son iguales o la otra donde ellas son diferentes. Cuando son del mismo tamaño se debe determinar un marco el cual será denotado con una variable k el cual nos indicará cuantos datos serán tomados en consideración para poder usar el modelo ARMA como se muestra en la Figura 2.2.

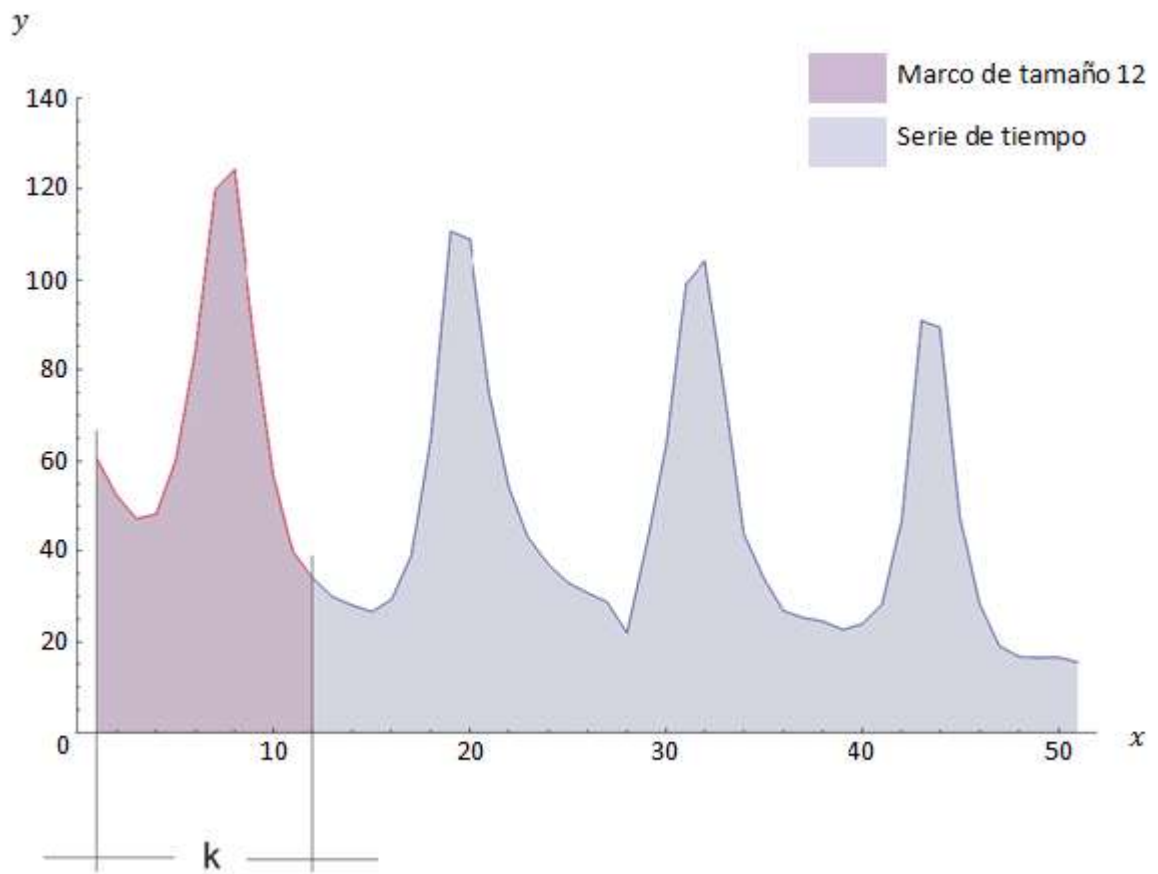


Figura 2.2 Serie de Tiempo con un Marco (k); el cual indica el número de valores de la gráfica que serán tomados en consideración.

Ahora se mostrará un ejemplo del funcionamiento del modelo ARMA. Consideremos 12 valores del marco (k), p y q del mismo tamaño. Los valores que tomarán las variables a_i serán 12 números aleatorios, al igual que b_i ; entonces p y q manejarán un marco de 12. Al principio, cuando se usa el modelo ARMA, no se consideran los errores ya que no existen; los errores se van generando mediante el avance del modelo y al igual que en el modelo autorregresivo se irán recorriendo los valores de $x_{t-1} \dots x_{t-p}$ y los de $\xi_{t-1} \dots \xi_{t-q}$, el valor de E es un número aleatorio. Todas estas variables son parte de la Ecuación (2.3).

En la Ecuación (2.4) se ve como no se toma en cuenta el error aún:

$$\hat{x}_t = a_1x_{t-1} + a_2x_{t-2} + a_3x_{t-3} + a_4x_{t-4} + a_5x_{t-5} + a_6x_{t-6} + a_7x_{t-7} + a_8x_{t-8} + a_9x_{t-9} + a_{10}x_{t-10} + a_{11}x_{t-11} + a_{12}x_{t-12} + E \quad (2.4)$$

En la Ecuación (2.5) se introduce el primer elemento del error:

$$\hat{x}_t = a_1x_{t-1} + a_2x_{t-2} + a_3x_{t-3} + a_4x_{t-4} + a_5x_{t-5} + a_6x_{t-6} + a_7x_{t-7} + a_8x_{t-8} + a_9x_{t-9} + a_{10}x_{t-10} + a_{11}x_{t-11} + a_{12}x_{t-12} + b_1\xi_{t-1} + E \quad (2.5)$$

En la Ecuación (2.6) se muestra el desarrollo ya avanzado del modelo ARMA cuando tiene 6 datos de error y este crecerá hasta llegar a 12, a y b no cambiarán solo se recorrerá los datos tanto de la serie de tiempo como la de los errores:

$$\hat{x}_t = a_1x_{t-1} + a_2x_{t-2} + a_3x_{t-3} + a_4x_{t-4} + a_5x_{t-5} + a_6x_{t-6} + a_7x_{t-7} + a_8x_{t-8} + a_9x_{t-9} + a_{10}x_{t-10} + a_{11}x_{t-11} + a_{12}x_{t-12} + b_1\xi_{t-1} + b_2\xi_{t-2} + b_3\xi_{t-3} + b_4\xi_{t-4} + b_5\xi_{t-5} + b_6\xi_{t-6} + E \quad (2.6)$$

En la Figura 2.3 se observa el avance del modelo ARMA después de haber predicho 6 valores, en esta figura vemos errores muy grandes en el método ARMA, por eso es necesario realizar varias pruebas.

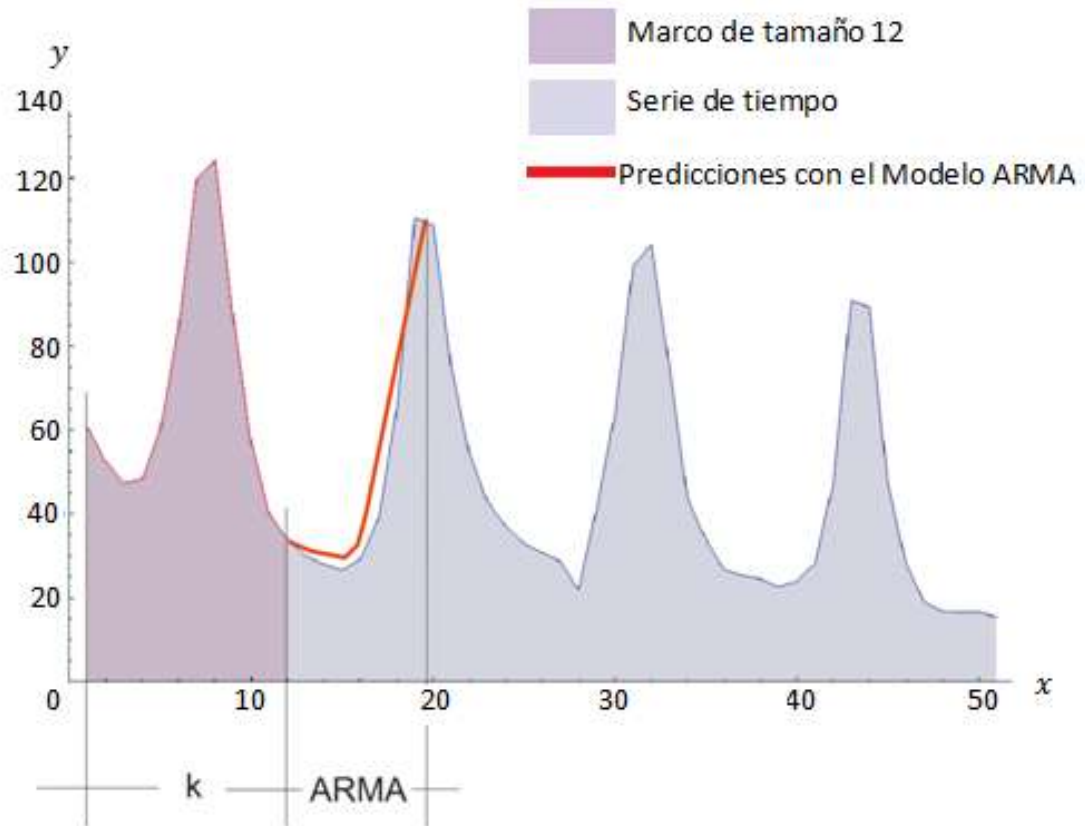


Figura 2.3 Serie de tiempo con seis valores calculados a través del modelo ARMA.

2.3 Algoritmos genéticos (AG)

Los Algoritmos Genéticos (AG) fueron introducidos en los 70's por J. H: Holland de la universidad de Michigan. Los AG forman parte de la Inteligencia Artificial la cual es una rama de Ciencias de la Computación [Wellin 2004]; son una Estrategia Evolutiva inspirada en la evolución biológica. A través de ellos podemos generar individuos que son sometidos a acciones aleatorias (como son la mutación y la recombinación genética). A partir de los individuos generados se realiza una selección para obtener el individuo más apto.

Un cromosoma es un conjunto de parámetros que definen una solución del problema que el algoritmo genético pretende resolver. Por lo general se representa como una serie de bits, números o cadenas las cuales están escritas en un lenguaje que debe ser definido antes de usarse para saber su representación.

La mutación es un operador genético, el cual consiste en una alteración o un cambio en la información genética del cromosoma, la mutación usa una probabilidad puntual de cambio decidiendo cuales cromosomas cambian y que porcentaje de ellos debe cambiar.

La recombinación es un operador genético, el cual consiste en una combinación genética de 2 cromosomas para obtener uno nuevo donde se decide cuantos elementos se intercambiarán sí para el nuevo cromosoma, como se ve en la Figura 2.4.

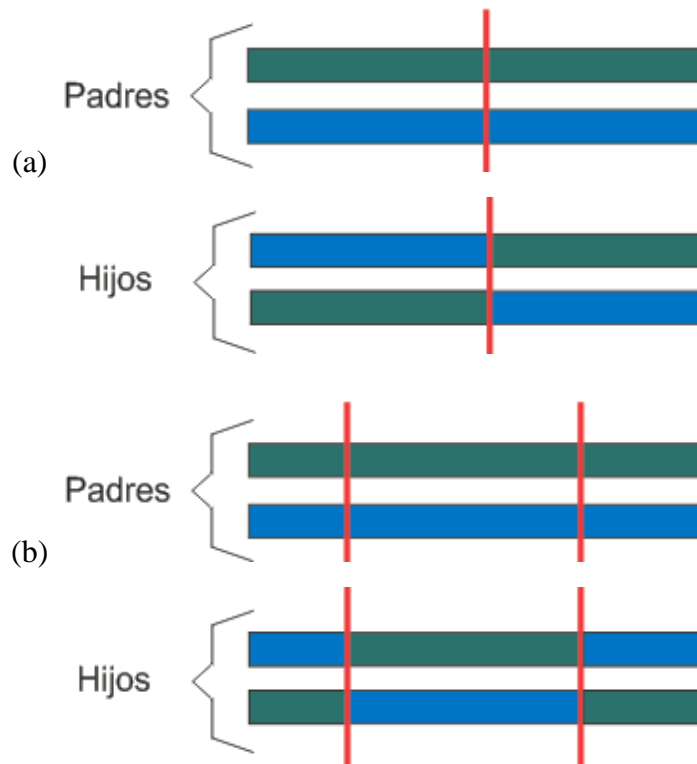


Figura 2.4 Ejemplos de diversos tipos de recombinación (a) combina la mitad de los padres para los nuevos hijos, (b) combina de diferentes lados para los padres y obtener hijos muy combinados.

En la Figura 2.4 se observa una raya roja que divide a los cromosomas y la cual indica que cantidad de cromosomas padres será proporcionado para generar a los nuevos cromosomas hijos, así como las diferentes posibilidades de hacer combinaciones de cromosomas en (a) los hijos tienen la mitad de cromosomas de cada padre mientras que en (b) se tienen diferentes partes de los cromosomas de los padres para generar a un cromosoma hijo.

Selección es una etapa muy importante en AG porque se tiene que escoger el individuo más apto siempre buscando un óptimo, en nuestro caso será siempre un máximo en la población de individuos como se ve en la Figura 2.5:

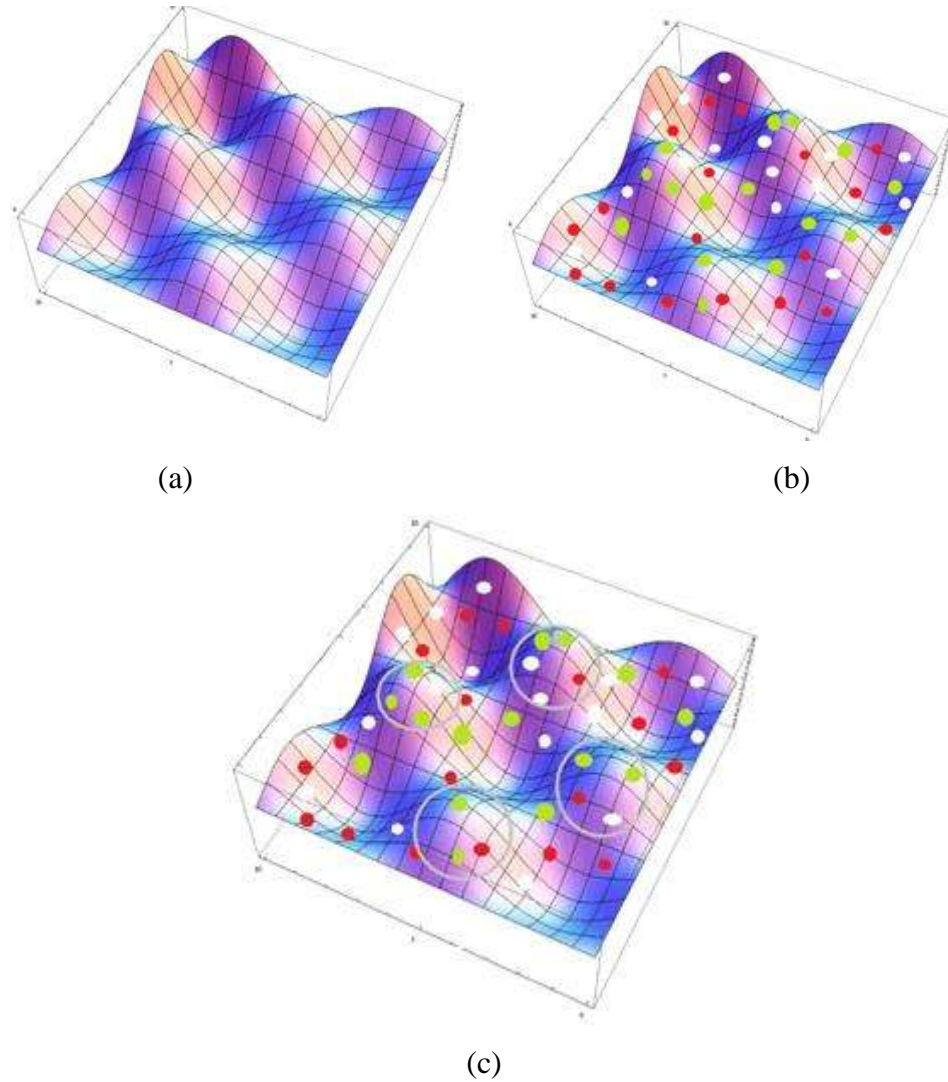


Figura 2.5 Selección (a) donde se situará la población, (b) se ve la población en diversas partes así como diversos individuos caracterizados por su color, (c) se selecciona a los individuos más aptos.

En la Figura 2.5 se observa que se escogieron los más cercanos a los picos máximos del lugar y estos serán los que llevarán a cabo operaciones de AG como mutación o recombinación. Estos individuos producirán las futuras generaciones y entonces se seguirá así continuamente hasta encontrar al mejor de todos.

Capítulo 3

Desarrollo de la programación de series de tiempo con Mathematica

Definir qué es lo que se quiere buscar es un buen comienzo para programar. En nuestro caso se van a crear cromosomas (constitución de elementos para formar un individuo; en algoritmos genéticos se refiere a datos los cuales pueden ser letras o números) los cuales serán los individuos a evaluar con la serie de tiempo original.

La función de aptitud dirá que tan bueno es un individuo para ser considerado apto para transmitir sus genes. En el modelo ARMA para saber si es un individuo es bueno debemos calcular un error y en base a ese error realizar la búsqueda.

La mutación y recombinación son funciones importantes de algoritmos genéticos puesto que son las funciones que se toman en cuenta para mejorar un cromosoma.

La serie de tiempo que se utilizará tiene las coordenadas de la ciudad de Morelia ubicándonos con una latitud de 19.69 y una longitud de -101.2 de la Universidad Michoacana de San Nicolás de Hidalgo.

Con Mathematica [Ruskeepää 2005] buscaremos la representación de la serie de tiempo de la ciudad de Morelia.

La función `WeatherData` de Mathematica tiene diversas opciones una usando la latitud y la longitud como se ve a continuación:

```
WeatherData[{latitud, longitud}]
```

La función regresara la ubicación representativa en Mathematica y para la ciudad de Morelia se debe aplicar de la manera siguiente:

```
In[102]:= WeatherData[{19.69, -101.2}]
```

```
Out[102]= MMMM
```


Otra de las opciones de la función WeatherData es poder obtener una lista de valores indicándole el lugar, la propiedad, las fechas de inicio, fin y por último si los datos serán dados por días, semanas, meses o años. Para el uso de la función se maneja así:

```
WeatherData["lugar", "propiedad", {{fecha de inicio}, {fecha de fin}, "tiempo" ]]
```

El tiempo representa si los datos que regresarán son en días, semanas, meses o años.

Ahora se especifica el lugar que será MMMM (ciudad de Morelia), la temperatura promedio del ambiente, se le indica la fecha de inicio de hace 10 años (1 de enero del 2000), la fecha de fin de la serie de tiempo será marcada hasta el 15 de octubre del 2010 y los datos serán mensuales los cuales están en el Anexo (pag. 43) Valores del Capítulo 3.

Para graficar estos valores utilizaremos la función DateListPlot de Mathematica; para usar el comando se debe poner la lista de datos; así como los valores que caracterizaran a las graficas.

```
DateListPlot[lista de datos, valores que caracterizan a la grafica]
```

Para este caso se ocupa decir que los puntos que se mostraran deben estar unidos entonces se utiliza Joined el cual debe ser verdadero). La gráfica resultante se muestra en la Figura 3.1.

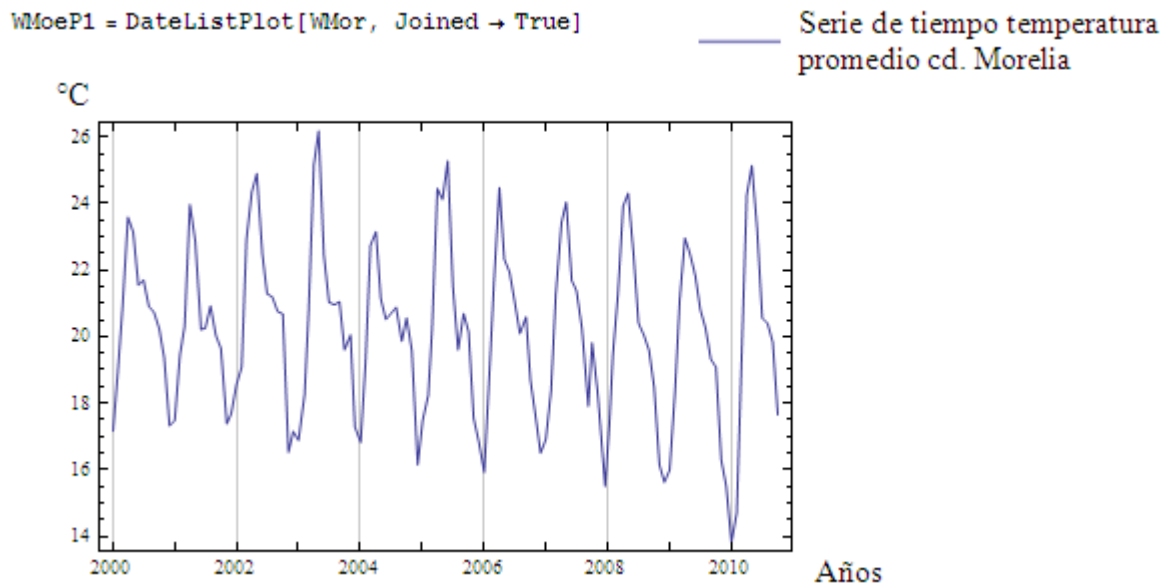


Figura 3.1 Gráfica de temperatura promedio mensual de la ciudad de Morelia, de 1° Enero del 2000 hasta el 15 de Octubre del 2010.

Con AG debemos determinar cómo crear el cromosoma en la ecuación (2.3) se observan los elementos que se forman parte del modelo ARMA; vemos que existen diferentes factores como a_i , b_i y E los cuales se manejarán como números aleatorios y que no dependen de los valores de la serie de tiempo, entonces podemos armar el cromosoma y que representen los parámetros a_i , b_i y E .

Para generar un cromosoma se utiliza la función `createChromosome` de `Evolvica` en la cual se le asigna un número de elementos que formarán al cromosoma así como el tipo de alfabeto que tendrá; la sintaxis es la siguiente:

```
createChromosome[número de elementos, tipo de alfabeto]
```

El cromosoma que se ve a continuación se muestra es un ejemplo de 25 valores aleatorios de -0.02 a 0.5, sus valores generan un modelo ARMA; el cual tendrá un error visible y una comparación de gráficas (con la serie de tiempo original y el modelo ARMA generado):

```
In[179]= chr1 = createChromosome[25,
      Alphabet -> {RandomReal[{-0.2, .5}]}]

Out[179]= chromo[q[undef], p[{1, {-0.0552139}},
      {2, {0.21028}}, {3, {0.165761}},
      {4, {-0.0027599}}, {5, {0.0230314}},
      {6, {0.302218}}, {7, {-0.0503097}},
      {8, {0.419762}}, {9, {0.183286}},
      {10, {0.464367}}, {11, {0.213319}},
      {12, {0.175732}}, {13, {0.277386}},
      {14, {0.209973}}, {15, {0.463261}},
      {16, {0.0879851}}, {17, {0.272469}},
      {18, {-0.172746}}, {19, {0.124714}},
      {20, {-0.00130838}}, {21, {0.37009}},
      {22, {-0.0246603}}, {23, {0.239503}},
      {24, {0.227965}}, {25, {0.342857}}]]
```

Ya se tiene el cromosoma ahora el cromosoma se separa en las variables de la Ecuación (2.3) a_i , b_i y E ; para llevar a cabo dicha operación se separa el cromosoma generado en el ejemplo anterior donde se indicara $E = c12$, $a_i = ai12$ y $b_i = bi12$; la manera de hacer esa separación es mediante corchetes dobles se coloca un valor numérico

en el cual están indicando el elemento que se desea obtener de un vector o lista; en ocasiones hay más de un valor la razón es porque se requiere obtener un valor específico como se ve en la Figura 3.2.

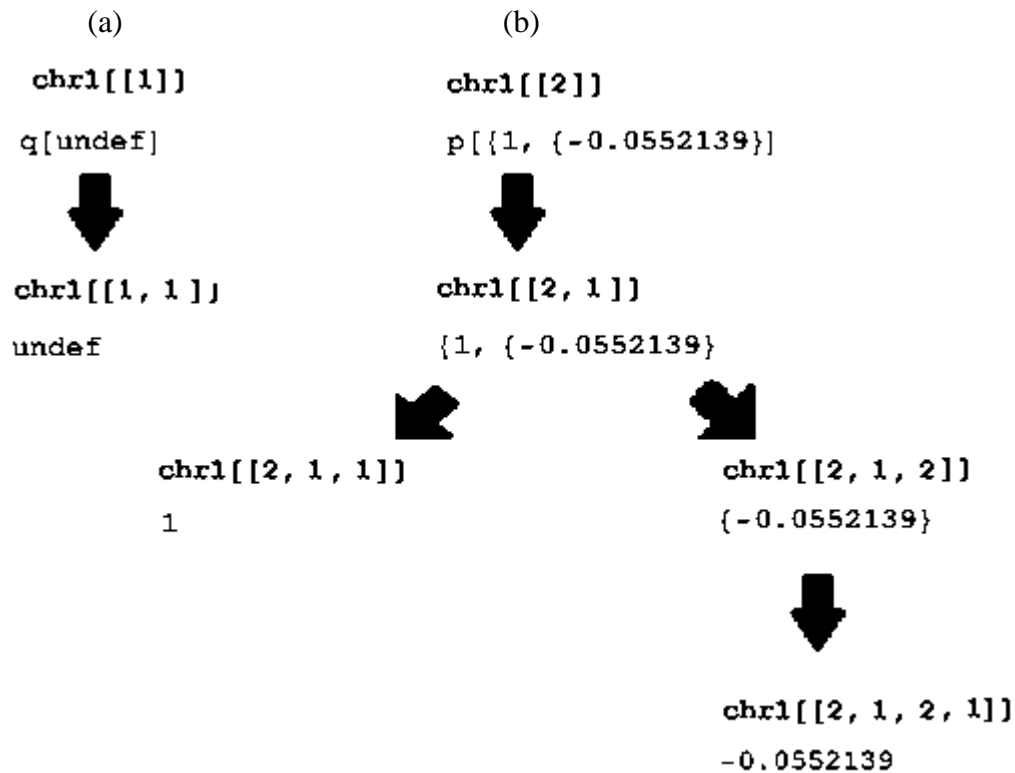


Figura 3.2 demostración del uso de corchetes dobles sobre un cromosoma (a) representa el primer elemento del chr1 mientras que (b) muestra el segundo elemento del chr1.

En la Figura 3.2 se observa la manera de obtener diferentes valores con (a) señala el valor de la calificación del cromosoma, en (b) muestra el elemento del cual requiriere separar el cromosoma en el vector a_i , b_i y el elemento E . Pero se necesita almacenar varios valores en los vectores a_i y b_i para realizar esta operación se requiere de la función Table de Mathematica la cual genera una lista de expresiones la sintaxis es la siguiente:

Table[expresión, {variable, valor mínimo, valor máximo}]

La variable maneja un valor mínimo hasta un valor máximo; donde la variable debe formar parte de la expresión que se esté manejando, ya que de otro modo solo se repetirá la expresión con el número de veces con respecto a la diferencia del valor máximo menos el

mínimo.

En la Figura 3.2 en (b) cuando `chr1[[2,1]]` regresa el valor de `{1, {-0.0552139}}` lo que significa que para obtener los otros elementos se debe cambiar el segundo elemento, el cual se maneja como la variable de la expresión para la función `Table`; para llevar a cabo la separación de los elementos que se necesitan se indica de la siguiente manera:

```
In[180]:= c12 = chr1[[2, 25, 2, 1]]
```

```
Out[180]= 0.342857
```

```
In[181]:= ai12 = Table[chr1[[2, i, 2, 1]], {i, 1, 12}]
```

```
Out[181]= {-0.0552139, 0.21028, 0.165761, -0.0027599,  
0.0230314, 0.302218, -0.0503097, 0.419762,  
0.183286, 0.464367, 0.213319, 0.175732}
```

```
In[182]:= bi12 = Table[chr1[[2, i, 2, 1]], {i, 13, 24}]
```

```
Out[182]= {0.277386, 0.209973, 0.463261, 0.0879851,  
0.272469, -0.172746, 0.124714, -0.00130838,  
0.37009, -0.0246603, 0.239503, 0.227965}
```

Ahora se extraen los valores de la serie de tiempo y para ello se utilizará también `Table` como se ve a continuación:

```
In[104]:= y = Table[WMor[[i, 2]], {i, Length[WMor]}]
```

```
Out[104]= {17.16, 19.08, 20.97, 23.59, 23.14, 21.55,  
21.7, 20.91, 20.73, 20.24, 19.35, 17.32, 17.48,  
19.43, 20.32, 23.97, 22.87, 20.21, 20.25,  
20.92, 20.03, 19.65, 17.38, 17.68, 18.57,  
19.08, 22.92, 24.35, 24.9, 22.49, 21.29, 21.19,  
20.76, 20.68, 16.52, 17.15, 16.88, 18.26, 20.77,  
25.16, 26.18, 22.46, 21.03, 20.96, 21.05, 19.6,  
20.05, 17.29, 16.82, 19.22, 22.74, 23.15, 21.13,  
20.52, 20.71, 20.88, 19.86, 20.56, 19.59, 16.15,  
17.49, 18.23, 20.23, 24.45, 24.12, 25.28, 21.53,  
19.6, 20.69, 20.14, 17.55, 16.83, 15.92, 18.96,  
21.65, 24.48, 22.33, 21.94, 21.03, 20.09, 20.59,  
18.69, 17.64, 16.5, 16.91, 18.38, 21.28, 23.42,  
24.04, 21.69, 21.38, 20.25, 17.91, 19.81, 18.25,  
15.5, 17.41, 19.44, 21.25, 23.92, 24.31, 22.44,  
20.4, 20.05, 19.59, 18.48, 16.15, 15.62, 15.99,  
18.31, 21.09, 22.96, 22.48, 21.83, 20.81, 20.25,  
19.32, 19.08, 16.3, 15.51, 13.82, 14.7, 19.24,  
24.25, 25.14, 23.24, 20.56, 20.4, 19.83, 17.65}
```

Una vez que se extraen los valores se utilizará el modelo ARMA implementando en una función que se desarrollo se llama f. Para el funcionamiento de f se ejemplificará con el siguiente algoritmo:

```

Procedimiento f(a, b, c, y)
j ← 1
k ← tamaño(a)
para i ← 1 hasta tamaño (a) hacer
    error(i) ← 0
fin para
para i ← 1 hasta tamaño (a) hacer
    yn(i) ← 0
fin para
para j ← 1 hasta tamaño(y) – k hacer
    ai(1) ← 0
    para i ← 1 hasta k hacer
        ai(i + 1) ← a(i) * y(i + j – 1)
    fin para
    bi(1) ← 0
    para i ← 1 hasta k hacer
        bi(i + 1) ← b(i) * error(i + j – 1)
    fin para
    yn(k + j) ← a(i + 1) + b(i + 1) + c
    error(k + j) ← y(k + j) – yn(k + 1)
fin para
fin Procedimiento

```

Esta función hace muchas iteraciones de la sumatoria del modelo ARMA, entre sus cálculos declara un vector inicial de ceros los cuales representan a los errores iniciales; así como los valores iniciales de la serie de tiempo que se está generando, después n elementos siguientes son calculados a través del modelo ARMA la cual necesita de un vector a_i , un

vector b_i , un número E que es representado como c en el algoritmo y los valores de y que representa a los valores de la serie de tiempo.

Entonces la función f regresa los valores del modelo ARMA la sintaxis es la siguiente:

$f[\text{vector } a, \text{vector } b, \text{numero } E, \text{valores de la serie de tiempo}]$

Ahora se cita el siguiente ejemplo donde se muestran valores del modelo ARMA con los elementos $ai12$, $bi12$, $c12$ que representa a E los datos obtenidos del cromosoma y los valores de la serie de tiempo original:

```
In[183]= yn12 = f[ai12, bi12, c12, y]
```

```
Out[183]= {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 45.6908,  
39.0516, 33.4685, 37.0468, 28.4852,  
34.4933, 34.2983, 37.8562, 32.401, 26.9645,  
19.5164, 16.5991, 19.9019, 22.8567,  
29.8236, 25.6062, 27.5098, 27.5333,  
29.1954, 35.4055, 36.1, 37.6767, 34.3556,  
31.0061, 28.0064, 27.7593, 27.8565, 26.8659,  
28.345, 26.9552, 26.4351, 26.7208, 25.9527,  
26.2569, 26.974, 30.1645, 30.5642, 31.1569,  
32.1316, 30.7902, 30.033, 30.4151, 32.1336,  
30.1961, 28.7547, 25.7255, 24.1062, 25.206,  
25.7666, 27.6805, 26.2207, 24.8265, 26.2933,  
29.9322, 33.1757, 35.2766, 35.036, 30.2554,  
28.2139, 26.3163, 27.3681, 30.5651, 29.1614,  
28.8501, 26.4217, 24.2845, 26.1001, 28.9742,  
29.0306, 28.6051, 27.1017, 25.5289, 26.368,  
28.4224, 28.9822, 29.664, 30.3424, 29.8862,  
30.1985, 29.8626, 29.2589, 26.9847, 26.4883,  
26.5787, 25.3802, 26.6553, 27.9226, 26.4389,  
27.5683, 30.6935, 29.8888, 30.9784, 31.3062,  
27.7703, 26.7059, 26.7636, 26.9457, 27.1747,  
27.043, 26.2268, 25.6743, 26.5616, 28.0893,  
29.3416, 28.1416, 27.0209, 26.2902, 25.9011,  
26.9616, 27.1339, 25.6425, 24.8004, 25.9521,  
27.8879, 30.2723, 32.1902, 29.9744, 0}
```

Ahora se hará la comparación de los valores originales de la serie de tiempo con el modelo ARMA generado por la función f como se ve en la Figura 3.3:

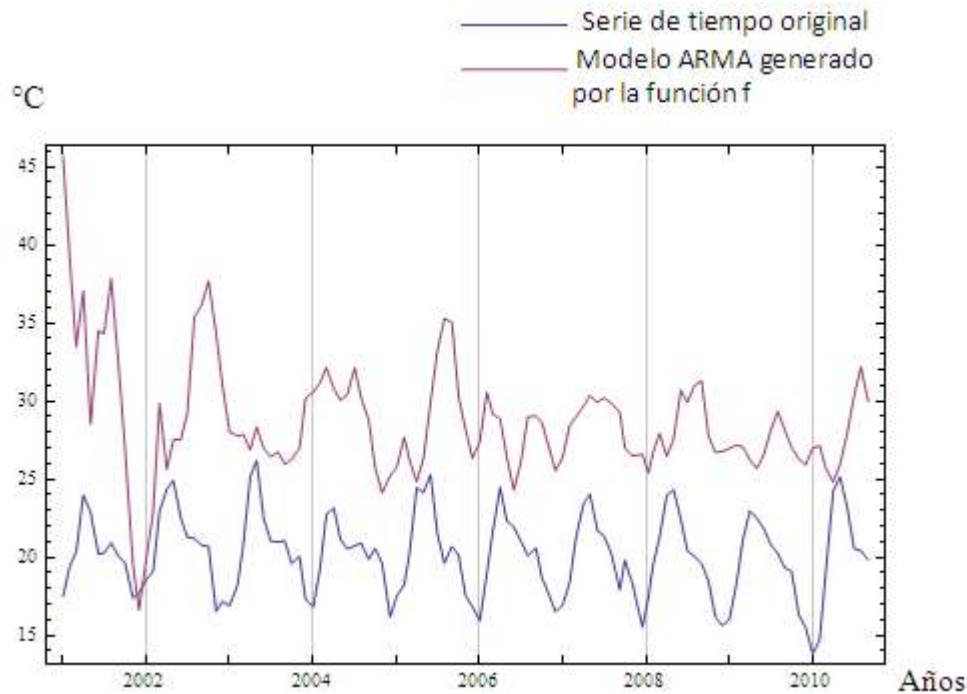


Figura 3.3 Comparación de Gráficas de temperatura promedio mensual de la ciudad de Morelia, original y con el modelo ARMA generado por la función f.

Probar un solo modelo ARMA generado de forma aleatoria no dio una buena solución, por tanto es necesaria la búsqueda de un modelo óptimo mediante el uso de AG.

La función *FitnessY* es otra función que se desarrollo en Mathematica para recibir como parámetros los valores de la serie de tiempo y el cromosoma que generara un modelo ARMA. Para el análisis de esta función que se creó; se realiza a través del siguiente algoritmo:

```

Procedimiento FitnessY(y, ca)
temp ← ca
c ← temp(1)
para i ← 2 hasta (tamaño (temp) + 1)/2 hacer

```

```

     $ai(i - 1) \leftarrow temp(i)$ 
  fin para
   $j \leftarrow 1$ 
  para  $i \leftarrow ((tamaños(temp) - 1)/2 + 2$  hasta  $tamaño(temp)$  hacer
     $bi(j) \leftarrow temp(i)$ 
     $j \leftarrow j + 1$ 
  fin para
   $yn \leftarrow f(ai, bi, c, y)$ 
   $j \leftarrow 1$ 
  para  $i \leftarrow tamaño(ai) + 1$  hasta  $tamaño(y) - 1$  hacer
     $nny(j) \leftarrow y(i)$ 
     $nyn(j) \leftarrow yn(i)$ 
     $j \leftarrow j + 1$ 
  fin para
   $aptitud \leftarrow 0$ 
  para  $i \leftarrow 1$  hasta  $tamaño(nny)$  hacer
     $aptitud \leftarrow aptitud + (nny(i) - nyn(i))^2$ 
  fin para
  si  $aptitud < Mejor$  entonces
     $Mejor \leftarrow aptitud$ 
    imprimir( $aptitud$ )
    imprimir( $ai$ )
    imprimir( $bi$ )
    imprimir( $c$ )
    graficar( $nny, nyn$ )
  fin si
  fin Procedimiento

```

La función FitnessY recibe dos parámetros, el valor de la serie de tiempo y el cromosoma con el que se va a evaluar. Esta función manda a llamar a f y así se obtiene el resultado de ARMA en la variable yn. Después la función calcula la suma de las diferencias

de dos listas de números enteros elevadas al cuadrado la cual representa el error de media cuadrática; en esa parte está almacenado el valor de aptitud; la cual deberá ser comparada con un margen de error que será asignado en una variable global Mejor con el siguiente valor:

```
In[191]:= Mejor = -1 * ListFit[y, yn12]
Out[191]= -9.06592 × 1014
```

Habría una comparación de aptitud es menor a la variable global Mejor, si es verdadero se hará un cambio de valor porque Mejor será igual al de aptitud. Cuando se esté utilizando AG se mostrara el valor de las mejores predicciones, así como sus modelos ARMA y graficas donde se muestra la comparación de la serie de tiempo con el Modelo ARMA.

El objetivo principal de este capítulo era mostrar una interpretación de la serie de tiempo en Mathematica así como una función adecuada para generar un modelo ARMA para realizar los futuros experimentos en AG en el capítulo siguiente.

Capítulo 4

Experimentos con Algoritmos Genéticos para Series de Tiempo

En el capítulo anterior se abordó en la programación necesaria para poder llevar a cabo los experimentos con AG en las series de tiempo (modelo ARMA).

Los AG sirven para resolver problemas de búsqueda y optimización. El Modelo ARMA es un método de predicción representado en la Ecuación (2.3). Para hacer la combinación entre el modelo ARMA y AG, y así poder encontrar el modelo ARMA que pueda acercarse a la temperatura de la ciudad de Morelia desde el 1 de enero 2000 hasta 15 de Octubre del 2010.

Para empezar los experimentos se define una función llamada GERMA representado en el siguiente algoritmo:

```
Procedimiento GERMA(k, y)  
    GApops ← Evolution(GA(numero de padres, numero de  
    combinaciones, operacion, numero de hijos, numero de generaciones,  
    tamaño del cromosoma, tipo de alfabeto, probabilidad de mutacion,  
    probabilidad de combinacion, funcion de evaluacion, poblacion inicial))  
fin Procedimiento
```

La cual contiene el número de marco y los valores de la serie de tiempo (los cuales fueron extraídos y almacenados en la variable *y*).

A la variable *GApops* se le asigna el valor de que se obtenga de la función *Evolution* que fue desarrollada para *Evolvica* [Jacob 1987]; esta variable se almacenara los individuos que en este caso son los modelos ARMA. La evolución que se lleva a cabo necesita de *GA* es la función de *Evolvica* en la cual se deben definir varios parámetros; entre ellos están:

- El número de padres son los elementos que serán llevados a la evolución.

- El número de combinaciones para este caso solo se ocupará 2.
- La función PLUS que representa el tipo de selección para producir la siguiente generación.
- El número de hijos son los elementos que deberán de tenerse y compararse para cada generación.
- El número de generaciones se refiere a las veces que se buscará el mejor de los individuos.
- El tamaño del cromosoma, el cual indica de qué tamaño es el vector.
- El tipo de alfabeto que se debe usar para los cromosomas.
- La probabilidad de realizar o no las operaciones de mutación y combinación.

Una vez que se han definido dichos parámetros se puede comenzar con los experimentos siempre que se tenga la aplicación de Evolva. En la Figura 4.1 se ve la serie de tiempo en la cual queremos encontrar la predicción del modelo ARMA y AG.

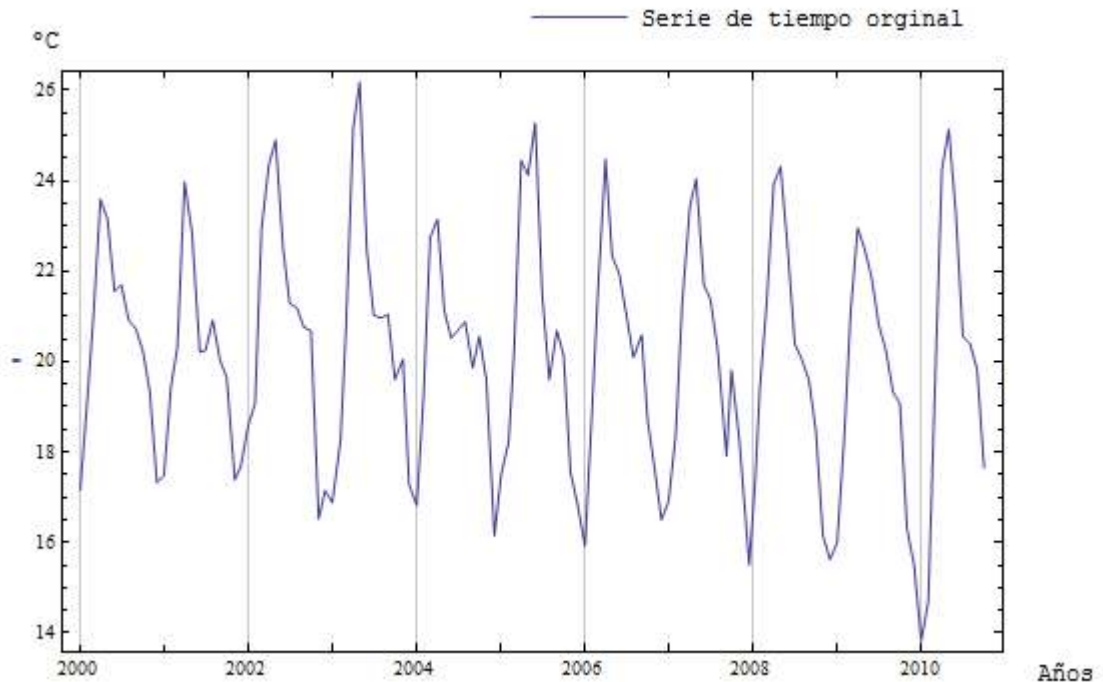


Figura 4.1 Serie de Tiempo del temperatura promedio mensual que hace en la ciudad de Morelia.

4.1 Experimentos con un marco de 12

Para el primer experimento ya se ha definido un tamaño de marco de 12 valores de la serie de tiempo mostrado en la Figura 4.2; los valores del marco se recorrerán mediante el modelo ARMA vaya avanzando hasta el último valor de la serie de tiempo:

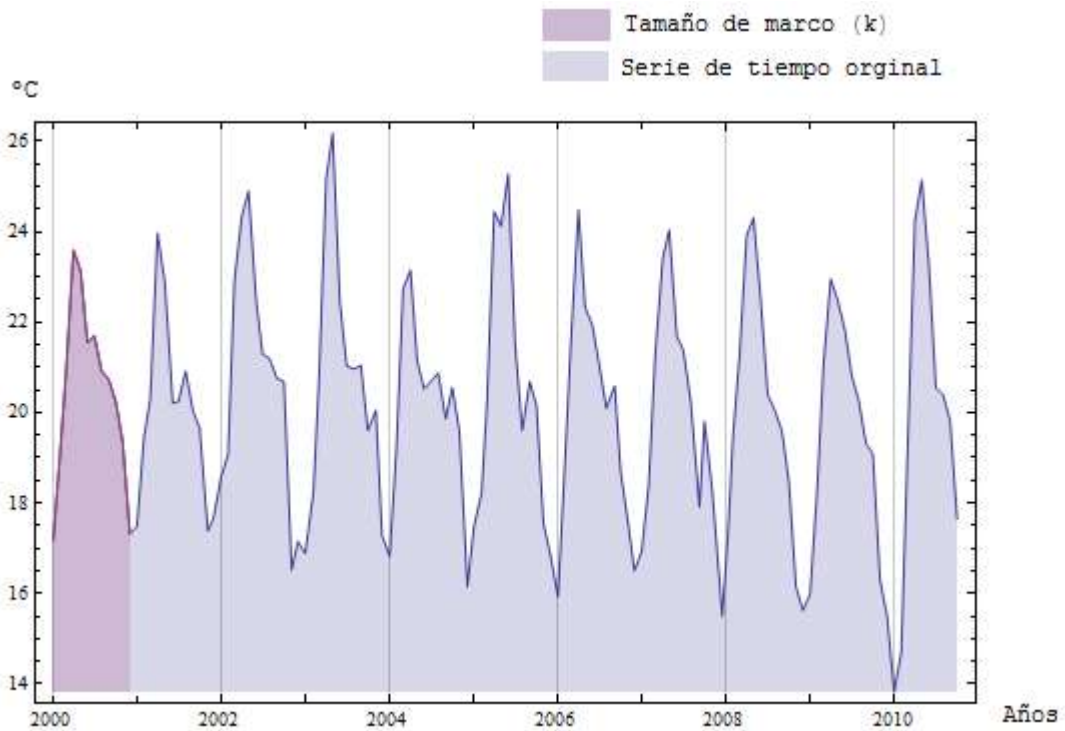
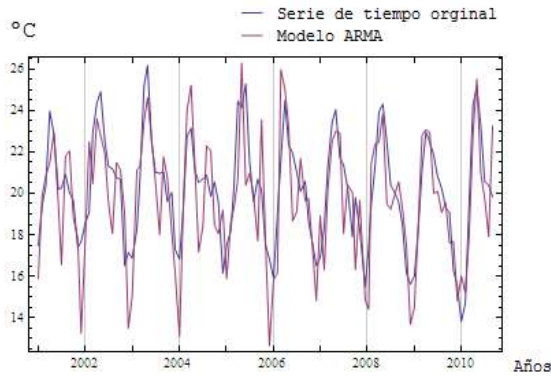


Figura 4.2 Serie de tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 12 valores del marco inicialmente.

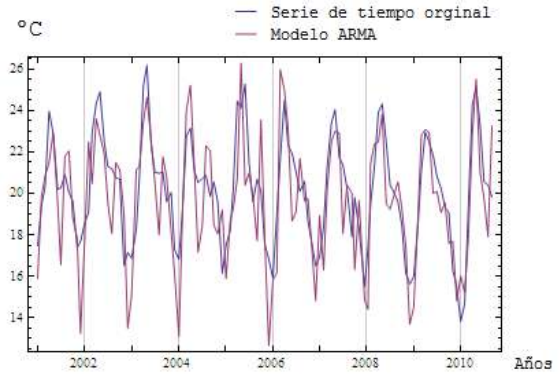
Después de haber definido el tamaño de marco se asignan los valores que hacen falta para el experimento. Entre ellos están el número de padres, el número de combinaciones, el número de hijos, las generaciones, el tamaño del cromosoma, el alfabeto y la función que se evaluara los resultados que se obtuvieron están mostrados en la Figura 4.3 y en la tabla 4.1.

Cuando se tiene ya todo definido solo queda hacer la prueba en Mathematica [Ruskeepää 2005] y Evolvica [Jacob 1987] como se muestra a continuación:

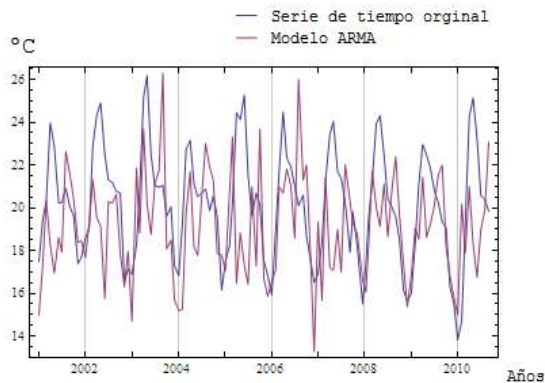
```
In[150]:= GEARMA[12, y]
          GA evolution ...
```



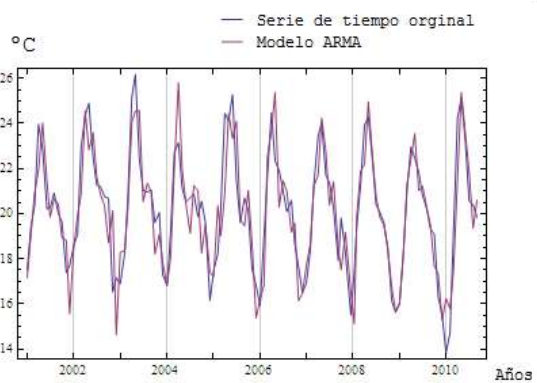
(a) Modelo ARMA el peor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 12 y un cromosoma de 25 valores.



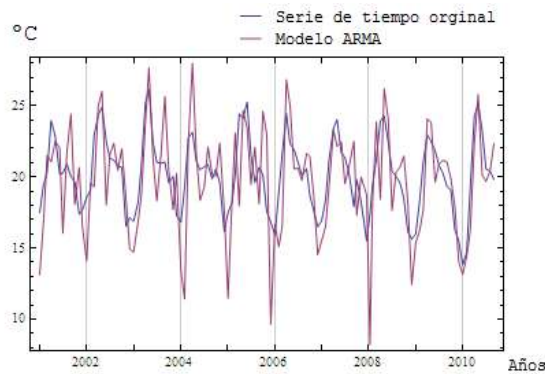
(b) Modelo ARMA el mejor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 12 y un cromosoma de 25 valores.



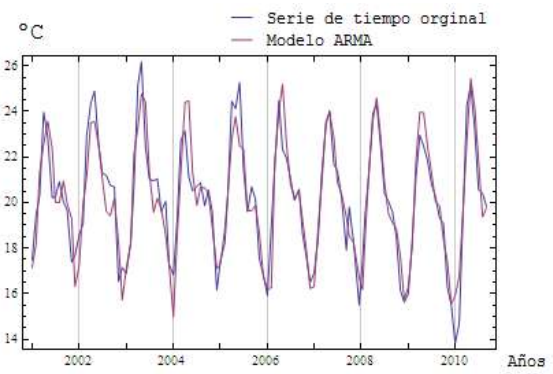
(c) Modelo ARMA el peor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 12 y un cromosoma de 25 valores.



(d) Modelo ARMA el mejor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 12 y un cromosoma de 25 valores.



(e) Modelo ARMA el peor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 12 y un cromosoma de 25 valores.



(f) Modelo ARMA el mejor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 12 y un cromosoma de 25 valores.

Figura 4.3 Resultados obtenidos con un tamaño de marco de 12 valores.

Tabla 4.1 Datos obtenidos de los experimentos con un tamaño de marco de 12 valores.

| Individuos | | Numero de | | | | Error de Media Cuadrática |
|------------|----------|-------------------|--------|-------|--------------|---------------------------|
| | | población inicial | padres | hijos | generaciones | |
| 400 | el peor | 400 | 400 | 200 | 200 | 407.347 |
| | | | | | | |
| | el mejor | 400 | 400 | 200 | 200 | 196.399 |
| | | | | | | |
| 800 | el peor | 800 | 800 | 400 | 400 | 981.364 |
| | | | | | | |
| | el mejor | 800 | 800 | 400 | 400 | 178.09 |
| | | | | | | |
| 2000 | el peor | 2000 | 2000 | 1000 | 1000 | 805.426 |
| | | | | | | |
| | el mejor | 2000 | 2000 | 1000 | 1000 | 121.198 |

Los experimentos realizados tienen un alfabeto de números reales aleatorios con un rango de -1 a 1 los cuales representan los 25 valores que contiene cada cromosoma de dichos experimentos. El cromosoma que se maneja se divide con los 12 primeros valores como el vector a_i , los otros 12 valores siguientes como b_i y el último valor representa a E variables de la ecuación (2.3).

La evolución de las generaciones se realiza con las operaciones de recombinación y mutación; con un 50% probabilidad de que cada elemento del cromosoma se cambie dependiendo la operación que se lleve a cabo.

En la Figura 4.3 y Tabla 4.1 tienen una relación en donde la figura (a) representa al peor individuo y (b) al mejor de una población inicial de 400 cromosomas, numero de padres de 400, número de hijos 200 y el numero de generaciones 200; en la figura (c) representa al peor individuo y (d) al mejor de una población inicial de 800 cromosomas, numero de padres de 800, número de hijos 400 y el numero de generaciones 400; en la figura (e) representa al peor individuo y (f) al mejor de una población inicial de 2000 cromosomas, numero de padres de 2000, número de hijos 1000 y el numero de generaciones 1000;

En los experimentos para determinar si un individuo es apto su función de aptitud está

basada en el error de media cuadrática y se determino que el mejor individuo de todos los resultados obtenidos se encuentra en la Figura 4.3 (f) y los datos que le corresponden al mejor individuo de la Tabla 4.1 con una población inicial de 2000 individuos, 1000 padres, 1000 hijos por cada generación, 1000 generaciones y un error de 121.198.

Por eso se considera el siguiente modelo ARMA en la Ecuación (4.1) y su representación generada en la Figura 4.4.

$$\begin{aligned} \hat{X}_t = & 0.319146x_{t-1} + 0.485414x_{t-2} - 0.0692122x_{t-3} - 0.142146x_{t-4} - 0.13638x_{t-5} \\ & - 0.0482947x_{t-6} + 0.211278x_{t-7} + 0.0837491x_{t-8} - 0.589367x_{t-9} + 0.0074352x_{t-10} \\ & + 0.420046x_{t-11} + 0.0559165x_{t-12} - 0.16417\xi_{t-1} - 0.08391\xi_{t-2} + 0.242992\xi_{t-3} \\ & - 0.0759278\xi_{t-4} - 0.178744\xi_{t-5} + 0.0873796\xi_{t-6} - 0.114532\xi_{t-7} - 0.019488\xi_{t-8} \\ & 0.237094\xi_{t-9} - 0.269312\xi_{t-10} - 0.262913\xi_{t-11} + 0.504433\xi_{t-12} + 0.669991 \end{aligned}$$

(4.1)

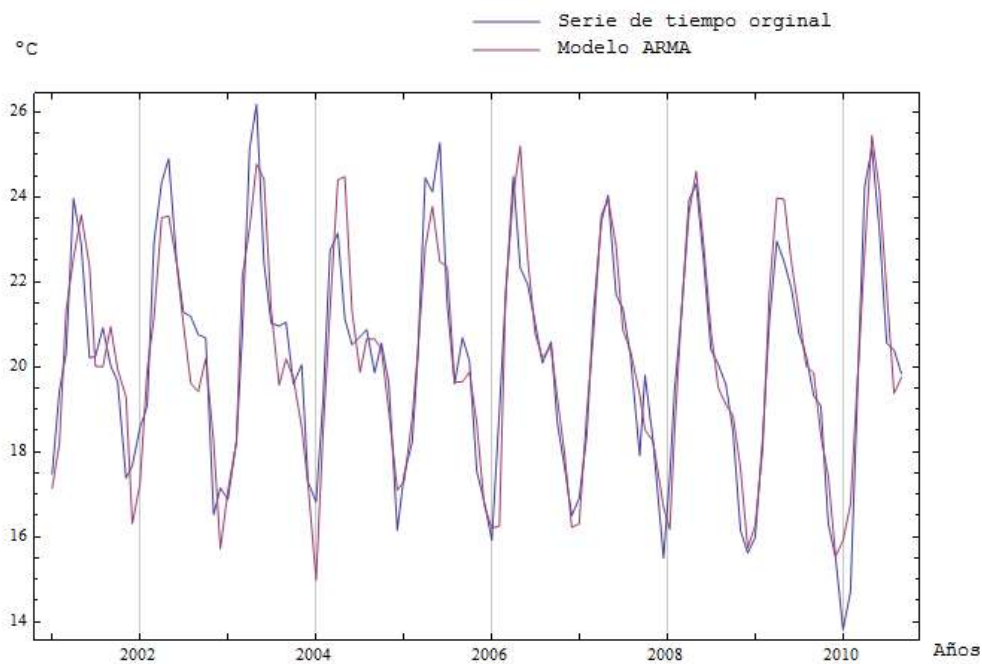


Figura 4.4 Representación la serie de tiempo original y el modelo ARMA de la Ecuación (4.1).

Los valores obtenidos de la Figura 4.4 están graficados entre el 01 de Enero del 2000 hasta el 01 de Septiembre del 2010.

Ahora en la Figura 4.5 basándose en las fechas en la que se encuentra graficada la Figura 4.4 se muestra una predicción de cuanto temperatura promedio se tendrá para el 01 de octubre del 2010.

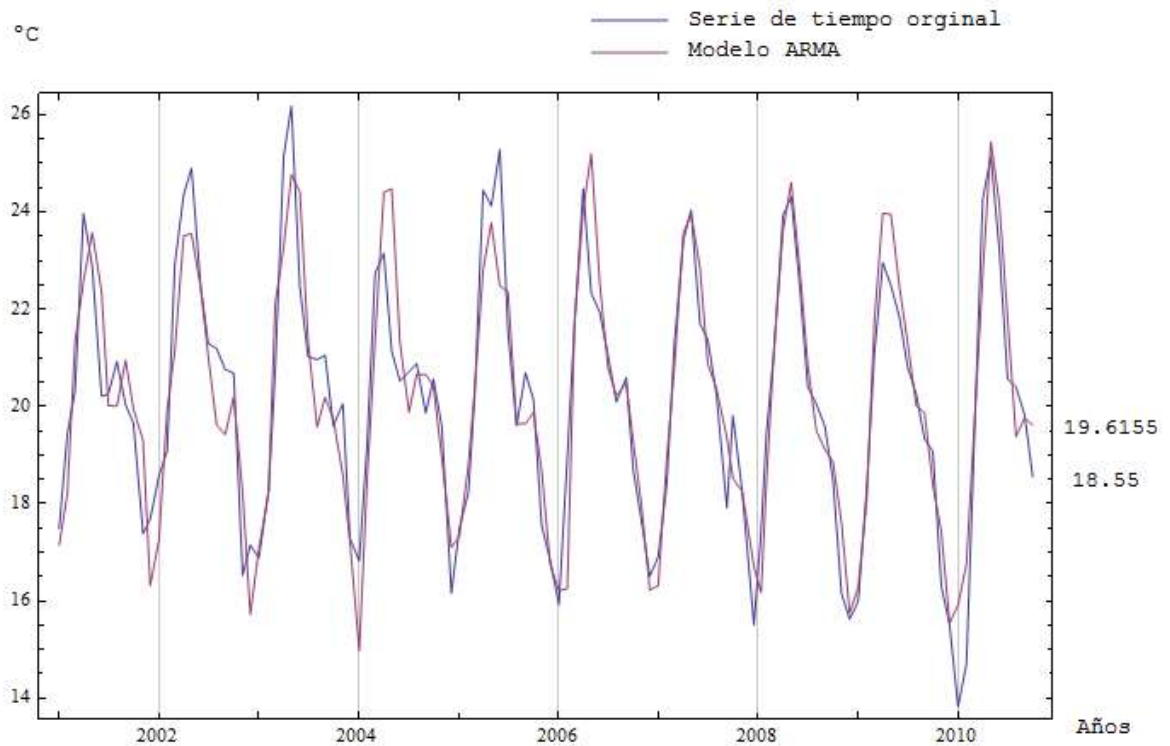


Figura 4.5 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 1.0655 °C.

En la Figura 4.5 el valor final que maneja la serie de tiempo es de 18.55°C mientras que el modelo ARMA calculo un valor de 19.6155°C, por lo tanto su error entre esos dos valores es de 1.0655°C.

4.2 Experimentos con un marco de 18

Para el segundo experimento ya se ha definido un marco de 18 valores de la serie de tiempo mostrado en la Figura 4.6; los valores del marco se recorrerán mediante el modelo ARMA vaya avanzando hasta el último valor de la serie de tiempo:

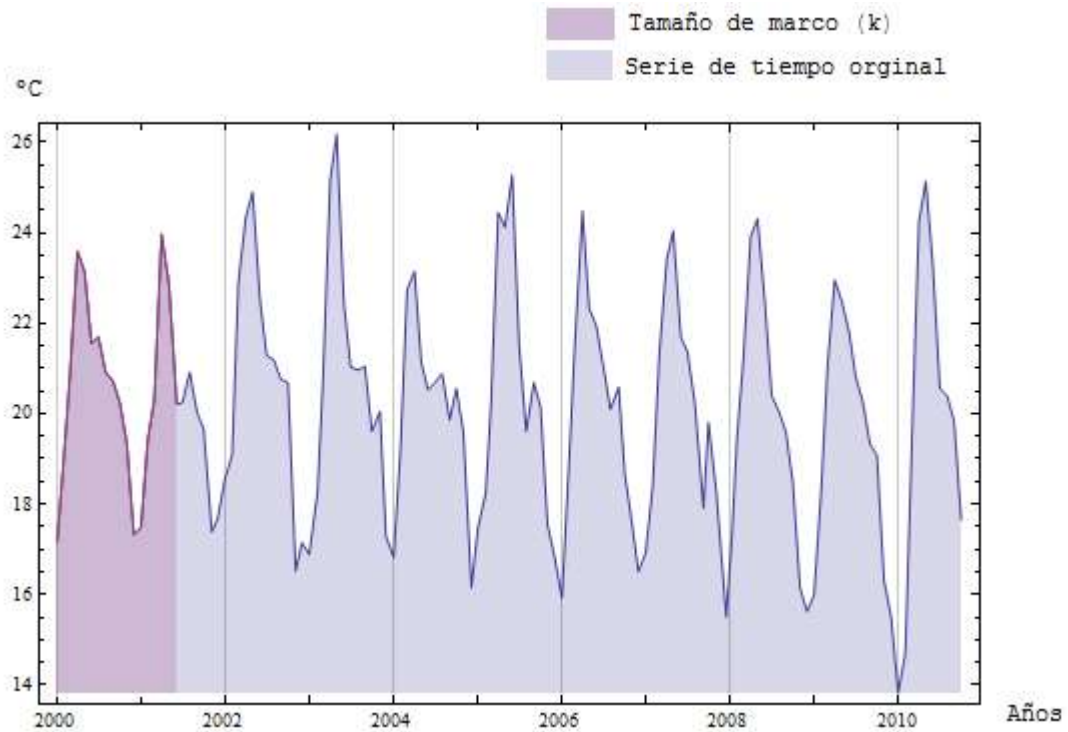
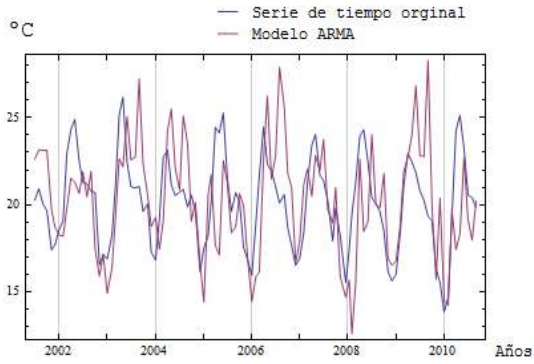


Figura 4.6 Serie de tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 18 valores del marco inicialmente.

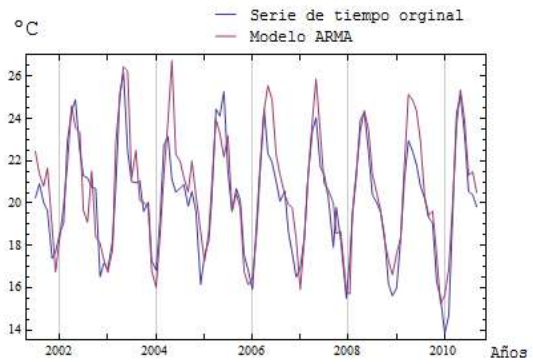
Después de haber definido el tamaño de marco debemos asignar los valores que hacen falta. Entre ellos están el número de padres, el número de combinaciones, el número de hijos, las generaciones, el tamaño del cromosoma, el alfabeto y la función que se evaluará los resultados que se obtuvieron están mostrados en la Figura 4.7 y en la tabla 4.2..

Cuando se tiene ya todo definido solo queda hacer la prueba en Mathematica y Evolvica como se muestra a continuación:

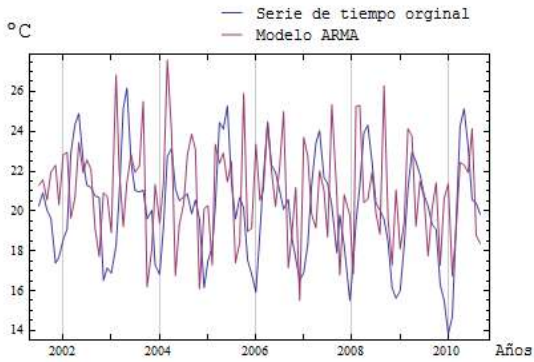
```
In[153]:= GEARMA[18, y]
          GA evolution ...
```



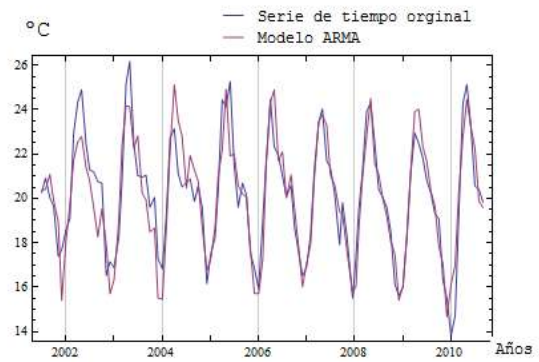
(a) Modelo ARMA el peor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 18 y un cromosoma de 37 valores.



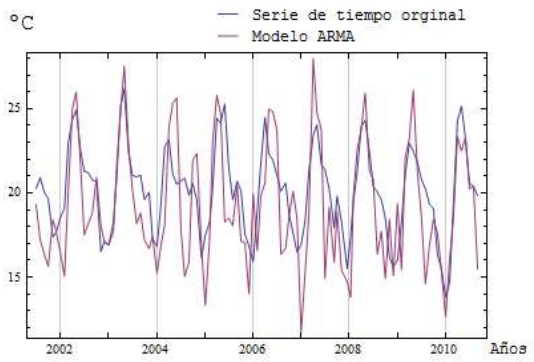
(b) Modelo ARMA el mejor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 18 y un cromosoma de 37 valores.



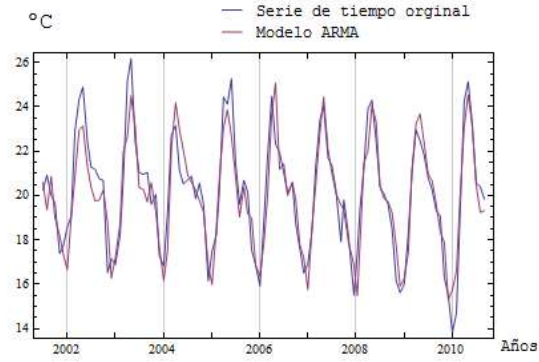
(c) Modelo ARMA el peor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 18 y un cromosoma de 37 valores.



(d) Modelo ARMA el mejor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 18 y un cromosoma de 37 valores.



(e) Modelo ARMA el peor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 18 y un cromosoma de 37 valores.



(f) Modelo ARMA el mejor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 18 y un cromosoma de 37 valores.

Figura 4.7 Resultados obtenidos con un tamaño de marco de 18 valores.

Tabla 4.2 Datos obtenidos de los experimentos con un tamaño de marco de 18 valores.

| Individuos | | Numero de | | | | Error de Media Cuadrática |
|------------|----------|-------------------|--------|-------|--------------|---------------------------|
| | | población inicial | padres | hijos | generaciones | |
| 400 | el peor | 400 | 400 | 200 | 200 | 915.729 |
| | | | | | | |
| | el mejor | 400 | 400 | 200 | 200 | 179.747 |
| | | | | | | |
| 800 | el peor | 800 | 800 | 400 | 400 | 983.3 |
| | | | | | | |
| | el mejor | 800 | 800 | 400 | 400 | 112.561 |
| | | | | | | |
| 2000 | el peor | 2000 | 2000 | 1000 | 1000 | 674.09 |
| | | | | | | |
| | el mejor | 2000 | 2000 | 1000 | 1000 | 96.0808 |

Los experimentos realizados tienen un alfabeto de números reales aleatorios con un rango de -0.5 a 0.5 los cuales representan los 37 valores que contiene cada cromosoma de dichos experimentos. El cromosoma que se maneja se divide con los 18 primeros valores como el vector a_i , los otros 18 valores siguientes como b_i y el último valor representa a E variables de la ecuación (2.3).

La evolución de las generaciones se realiza con las operaciones de recombinación y mutación; con un 50% probabilidad de que cada elemento del cromosoma se cambie dependiendo la operación que se lleve a cabo.

En la Figura 4.7 y Tabla 4.2 tienen una relación en donde la figura (a) representa al peor individuo y (b) al mejor de una población inicial de 400 cromosomas, numero de padres de 400, número de hijos 200 y el numero de generaciones 200; en la figura (c) representa al peor individuo y (d) al mejor de una población inicial de 800 cromosomas, numero de padres de 800, número de hijos 400 y el numero de generaciones 400; en la figura (e) representa al peor individuo y (f) al mejor de una población inicial de 2000 cromosomas, numero de padres de 2000, número de hijos 1000 y el numero de generaciones 1000;

En los experimentos para determinar si un individuo es apto su función de aptitud está

basada en el error de media cuadrática y se determino que el mejor individuo de todos los resultados obtenidos se encuentra en la Figura 4.7 (f) y los datos que le corresponden al mejor individuo de la Tabla 4.2 con una población inicial de 2000 individuos, 2000 padres, 1000 hijos por cada generación, 1000 generaciones y un error de 96.0808.

Por eso se considera el siguiente modelo ARMA en la Ecuación (4.2) y su representación generada en la Figura 4.8.

$$\begin{aligned}
 \hat{x}_t = & 0.226324x_{t-1} - 0.0829998x_{t-2} - 0.410802x_{t-3} - 0.168129x_{t-4} \\
 & -0.0620341x_{t-5} + 0.291494x_{t-6} + 0.301808x_{t-7} + 0.188712x_{t-8} \\
 & +0.189588x_{t-9} - 0.342541x_{t-10} + 0.115436x_{t-11} + 0.00677921x_{t-12} \\
 & -0.265021x_{t-13} + 0.151239x_{t-14} + 0.171055x_{t-15} + 0.0326292x_{t-16} \\
 & +0.191599x_{t-17} + 0.154842x_{t-1} - 0.133132\xi_{t-1} + 0.0841412\xi_{t-2} \\
 & +0.320896\xi_{t-3} + 0.145327\xi_{t-4} + 0.10131\xi_{t-5} - 0.124643\xi_{t-6} \\
 & -0.232867\xi_{t-7} - 0.0485142\xi_{t-8} - 0.168442\xi_{t-9} - 0.0347197\xi_{t-10} \\
 & -0.0434155\xi_{t-11} + 0.279356\xi_{t-12} + 0.364684\xi_{t-13} - 0.110165\xi_{t-14} \\
 & -0.337169\xi_{t-15} - 0.228453\xi_{t-16} - 0.289894\xi_{t-17} \\
 & +0.222659\xi_{t-18} + 0.339734
 \end{aligned}
 \tag{4.2}$$

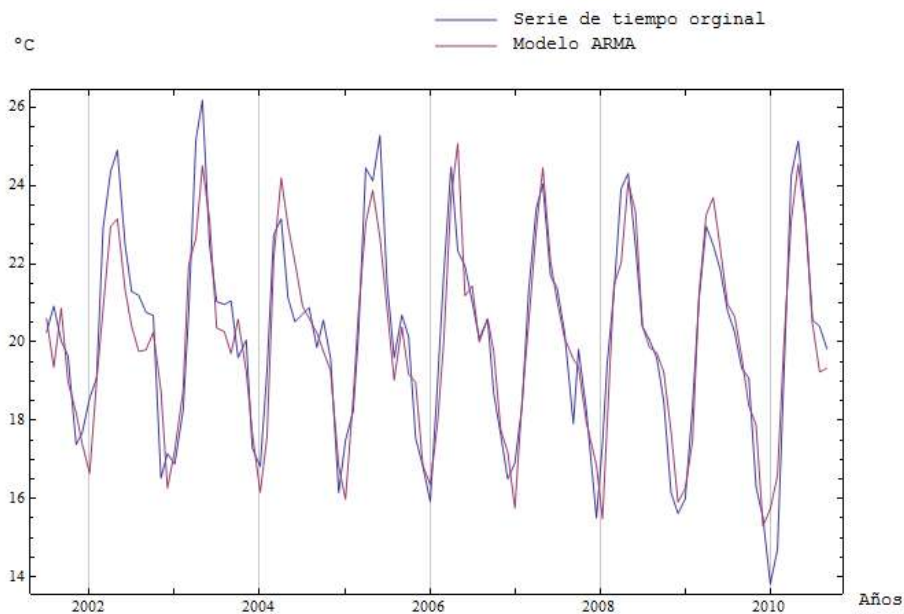


Figura 4.8 Representación la serie de tiempo original y el modelo ARMA de la Ecuación (4.2).

Los valores obtenidos de la Figura 4.8 están graficados entre el 01 de Enero del 2000 hasta el 01 de Septiembre del 2010.

Ahora en la Figura 4.9 basándose en las fechas en la que se encuentra graficada la Figura 4.8 se muestra una predicción de cuanta temperatura promedio se tendrá para el 01 de octubre del 2010.

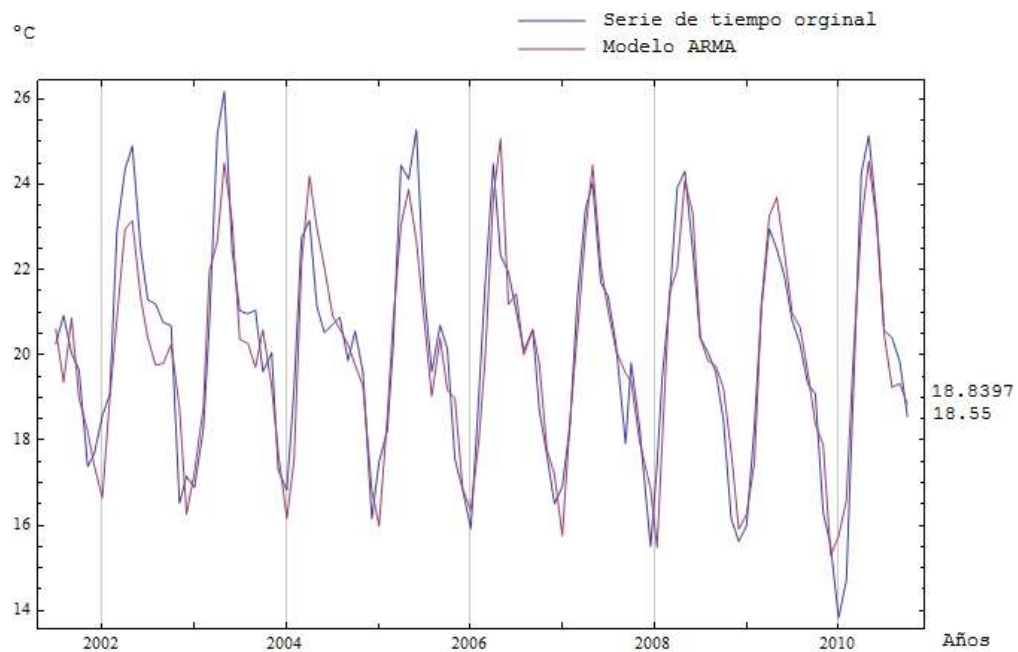


Figura 4.9 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 0.2897 °C.

En la Figura 4.9 el valor final que maneja la serie de tiempo es de 18.55°C mientras que el modelo ARMA calculo un valor de 18.8397°C, por lo tanto su error entre esos dos valores es de 0.2897°C.

4.3 Experimentos con un marco de 24

Para el segundo experimento ya se ha definido un marco de 24 valores de la serie de tiempo mostrado en la Figura 4.10; los valores del marco se recorrerán mediante el modelo ARMA vaya avanzando hasta el último valor de la serie de tiempo:

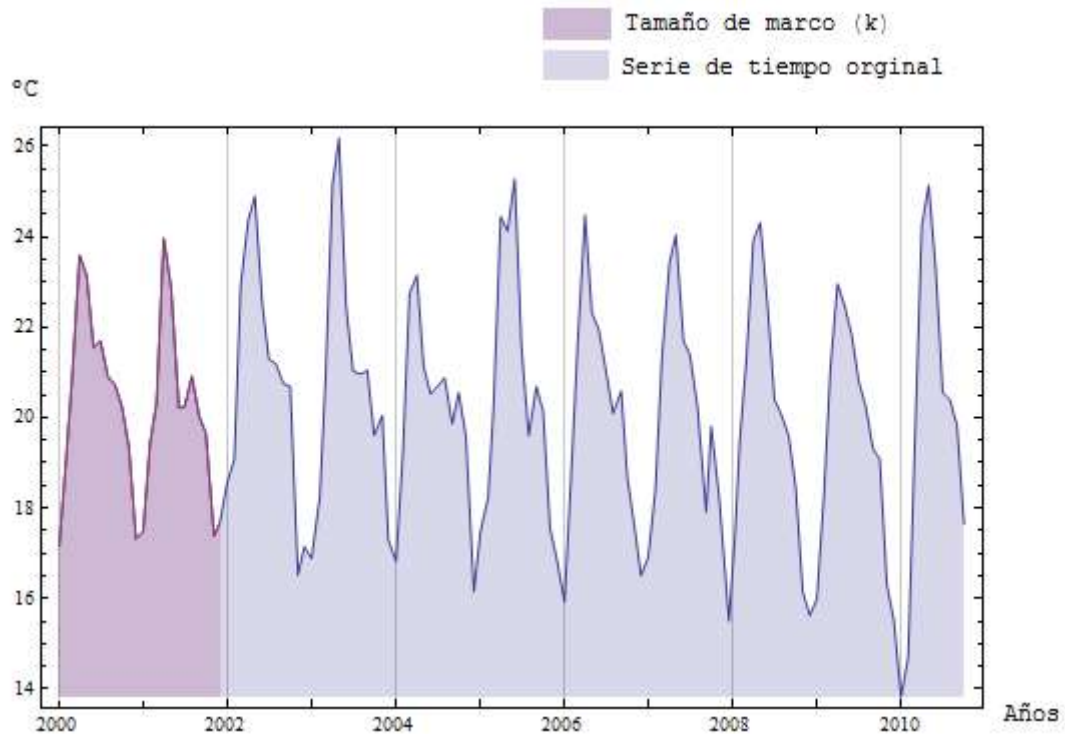
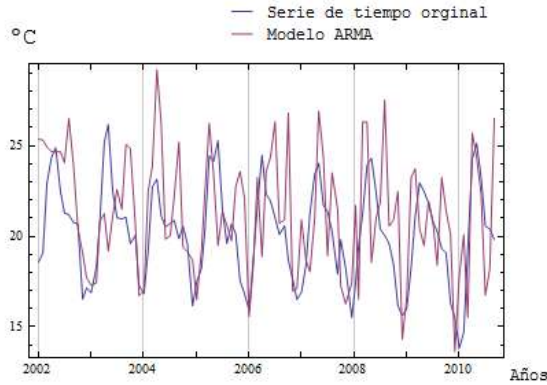


Figura 4.10 Serie de Tiempo de temperatura promedio de Morelia donde se muestra de color más sombreado sobre la parte representativa de los primeros 24 valores del marco inicialmente.

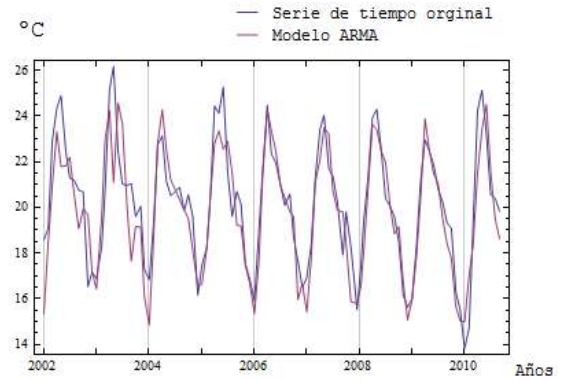
Después de haber definido el tamaño de marco debemos asignar los valores que hacen falta. Entre ellos están el número de padres, el número de combinaciones, el número de hijos, las generaciones, el tamaño del cromosoma, el alfabeto y la función que se evaluará los resultados que se obtuvieron están mostrados en la Figura 4.11 y en la tabla 4.3..

Cuando se tiene ya todo definido solo queda hacer la prueba en Mathematica y Evolvica como se muestra a continuación:

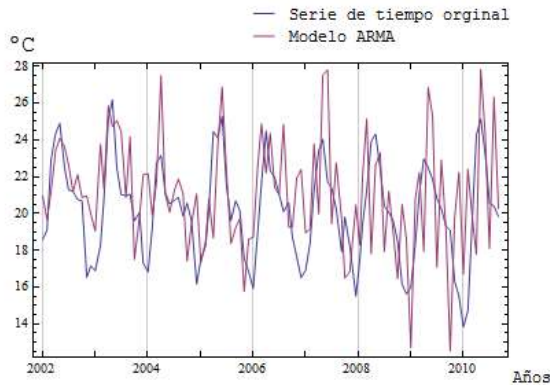
```
In[154]:= GEARMA[24, y]
          GA evolution ...
```



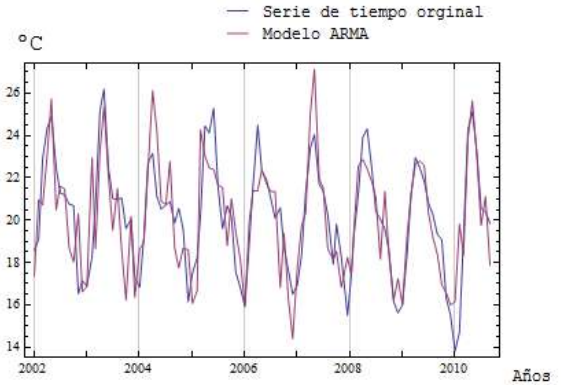
(a) Modelo ARMA el peor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 24 y un cromosoma de 49 valores.



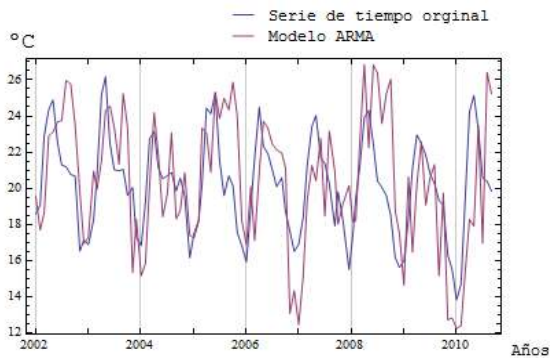
(b) Modelo ARMA el mejor individuo con una población de 400, las generaciones de 200, un tamaño de marco de 24 y un cromosoma de 49 valores.



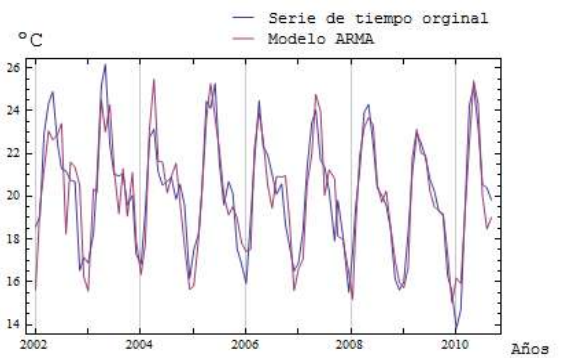
(c) Modelo ARMA el peor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 24 y un cromosoma de 49 valores.



(d) Modelo ARMA el mejor individuo con una población de 800, las generaciones de 400, un tamaño de marco de 24 y un cromosoma de 49 valores.



(e) Modelo ARMA el peor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 24 y un cromosoma de 49 valores.



(f) Modelo ARMA el mejor individuo con una población de 2000, las generaciones de 1000, un tamaño de marco de 24 y un cromosoma de 49 valores.

Figura 4.11 Resultados obtenidos con un tamaño de marco de 24 valores.

Tabla 4.3 Datos obtenidos de los experimentos con un tamaño de marco de 24 valores.

| Individuos | | Numero de | | | | Error de Media Cuadrática |
|------------|----------|-------------------|--------|-------|--------------|---------------------------|
| | | población inicial | padres | hijos | generaciones | |
| 400 | el peor | 400 | 400 | 200 | 200 | 876.528 |
| | | | | | | |
| | el mejor | 400 | 400 | 200 | 200 | 94.3926 |
| | | | | | | |
| 800 | el peor | 800 | 800 | 400 | 400 | 835.843 |
| | | | | | | |
| | el mejor | 800 | 800 | 400 | 400 | 215.056 |
| | | | | | | |
| 2000 | el peor | 2000 | 2000 | 1000 | 1000 | 828.359 |
| | | | | | | |
| | el mejor | 2000 | 2000 | 1000 | 1000 | 104.759 |

Los experimentos realizados tienen un alfabeto de números reales aleatorios con un rango de -0.5 a 0.5 los cuales representan los 49 valores que contiene cada cromosoma de dichos experimentos. El cromosoma que se maneja se divide con los 24 primeros valores como el vector a_i , los otros 24 valores siguientes como b_i y el último valor representa a E variables de la ecuación (2.3).

La evolución de las generaciones se realiza con las operaciones de recombinación y mutación; con un 50% probabilidad de que cada elemento del cromosoma se cambie dependiendo la operación que se lleve a cabo.

En la Figura 4.11 y Tabla 4.3 tienen una relación en donde la figura (a) representa al peor individuo y (b) al mejor de una población inicial de 400 cromosomas, número de padres de 400, número de hijos 200 y el número de generaciones 200; en la figura (c) representa al peor individuo y (d) al mejor de una población inicial de 800 cromosomas, número de padres de 800, número de hijos 400 y el número de generaciones 400; en la figura (e) representa al peor individuo y (f) al mejor de una población inicial de 2000 cromosomas, número de padres de 2000, número de hijos 1000 y el número de generaciones 1000;

En los experimentos para determinar si un individuo es apto su función de aptitud está

basada en el error de media cuadrática y se determino que el mejor individuo de todos los resultados obtenidos se encuentra en la Figura 4.11 (f) y los datos que le corresponden al mejor individuo de la Tabla 4.3 con una población inicial de 400 individuos, 400 padres, 200 hijos por cada generación, 200 generaciones y un error de 94.3962.

Por eso se considera el siguiente modelo ARMA en la Ecuación (4.3) y su representación generada en la Figura 4.12.

$$\begin{aligned} \hat{x}_t = & 0.473625x_{t-1} + 0.159361x_{t-2} - 0.12959x_{t-3} + 0.00783119x_{t-4} - 0.0536532x_{t-5} \\ & - 0.00954409x_{t-6} - 0.0380638x_{t-7} - 0.199913x_{t-8} + 0.101656x_{t-9} + 0.053676x_{t-10} \\ & + 0.045651x_{t-11} + 0.043821x_{t-12} - 0.0355498x_{t-13} + 0.177403x_{t-14} + 0.03656x_{t-15} \\ & - 0.0708132x_{t-16} - 0.16284x_{t-17} + 0.0607093x_{t-18} + 0.21161x_{t-19} + 0.282595x_{t-20} \\ & + 0.0019126x_{t-21} + 0.128026x_{t-22} - 0.193019x_{t-23} + 0.06319x_{t-24} + 0.488819\xi_{t-1} \\ & + 0.403024\xi_{t-2} - 0.0248942\xi_{t-3} + 0.421967\xi_{t-4} - 0.184938\xi_{t-5} + 0.29457\xi_{t-6} \\ & + 0.333374\xi_{t-7} + 0.179898\xi_{t-8} + 0.211833\xi_{t-9} + 0.254607\xi_{t-10} + 0.2462\xi_{t-11} \\ & + 0.283918\xi_{t-12} + 0.416219\xi_{t-13} + 0.248942\xi_{t-14} + 0.287744\xi_{t-15} + 0.435725\xi_{t-16} \\ & + 0.433586\xi_{t-17} - 0.0608951\xi_{t-18} + 0.306672\xi_{t-19} - 0.319774\xi_{t-20} - 0.11965\xi_{t-21} \\ & + 0.221835\xi_{t-22} + 0.340494\xi_{t-23} + 0.152315\xi_{t-24} + 0.326149 \end{aligned}$$

(4.3)

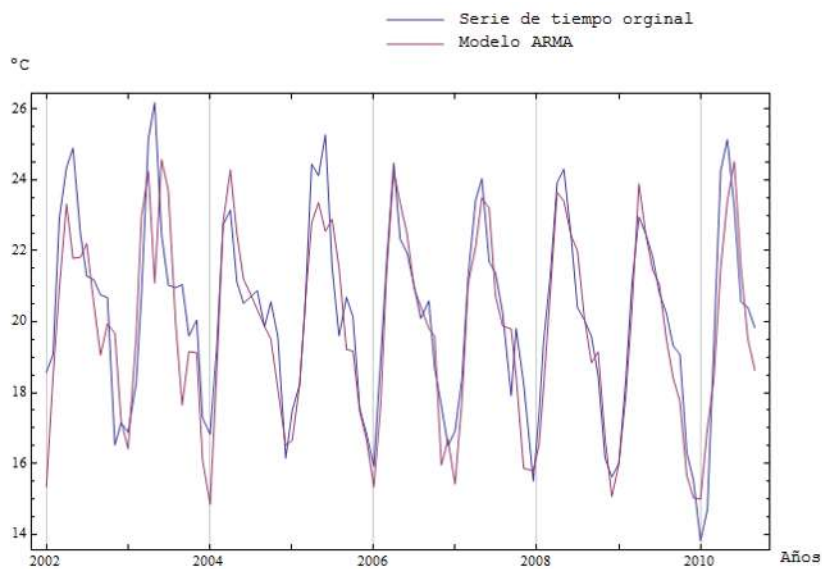


Figura 4.12 Representación la serie de tiempo original y el modelo ARMA de la Ecuación (4.4).

Los valores obtenidos de la Figura 4.12 están graficados entre el 01 de Enero del 2000 hasta el 01 de Septiembre del 2010.

Ahora en la Figura 4.13 basándose en las fechas en la que se encuentra graficada la Figura 4.12 se muestra una predicción de cuanto temperatura promedio se tendrá para el 01 de octubre del 2010.

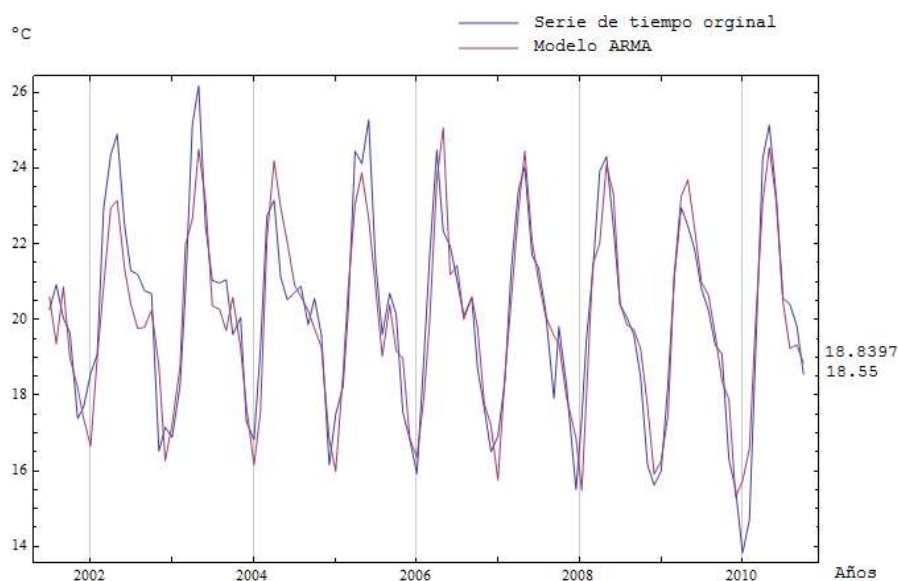


Figura 4.13 Predicción del modelo ARMA en la serie de tiempo de la temperatura promedio con un error de diferencia de 0.2325 °C.

En la Figura 4.8 el valor final que maneja la serie de tiempo es de 18.55°C mientras que el modelo ARMA calculo un valor de 18.7828°C, por lo tanto su error entre esos dos valores es de 0.2325°C.

En los experimentos realizados se observan los múltiples cambios que se presentaban con distintos marcos y con distintos números de individuos. Los resultados obtenidos muestran que el marco de 24 es el mejor con 400 de población inicial, 400 padres, 200 hijos y 200 generaciones. Sin embargo falta considerar que manejar un tamaño de marco de 24 y generando un individuo 49 valores se hubiera esperado que con una población más grande se obtuviera un mejor resultado más óptimo.

Capítulo 5

Conclusiones

Como se ha visto en el capítulo anterior se mostraron los resultados de series de tiempo con el modelo ARMA a través de Algoritmos Genéticos. En estos resultados se analiza el peor y el mejor individuo de cada uno de los experimentos, donde se regresa un error de media cuadrática el cual representa la función de aptitud. Se muestra también los coeficientes de la ecuación del modelo ARMA de los mejores individuos de cada experimento.

Con la serie de tiempo que se estuvo trabajando durante esta tesis fue de la temperatura promedio mensual de la ciudad de Morelia la cual se maneja en °C. Los resultados obtenidos variaron con respecto al tamaño de marco 12, 18 y 24; en dichos cambios se observó cómo variaban los valores entre los distintas pruebas de 400, 800 y 2000 individuos con cada marco que se uso.

La razón por la variación del alfabeto diferente entre los marcos es por que cuando se manejan mas valores se requiere de una población o generación mayor, entonces se trata de mostrar la diferencia entre diferentes marcos, pero con la misma población inicial, padres, hijos y generaciones; por esa razón se maneja diferentes alfabetos en el marco de 12 se uso uno de -1 a 1 y los marcos de 18 y 24 es de -0.5 a 0.5.

Puede variar el resultado óptimo como se ve en las tablas de los experimentos donde más varia es el experimento 3er experimento donde el mejor individuo de una población inicial de 400 fue mas óptimo que el mejor de una población inicial de 2000. Esto se debe al uso de numeros aleatorios reales por lo cual los resultados óptimos varían, pero al final dan un resultado óptimo y este puede ser a la mitad del proceso o incluso hasta el final de los AG's.

Las series de tiempo han servido durante muchos años para poder saber como está una empresa. Por ejemplo, si se tuviera una tienda y se quiere saber que venderá en un día de la

semana, como el lunes no se podría comparar con lo que vendió un día o dos días antes, sino deberá tomarse en consideración el lunes pasado; así como también podría fijarse si los lunes de semanas atrás se vendió igual y en caso de que no que factores intervinieron para que bajara o aumentara las ventas.

Entonces aquí en esta tesis se muestran las predicciones posibles de la serie de tiempo del clima de Morelia donde este valor nos dirá que temperatura podrá hacer. Este valor ya nos da una oportunidad de determinar si conviene o no realizar ciertas acciones.

El método que se empleó en esta tesis solo analiza la serie de tiempo de Morelia. Sin embargo, el método es factible para cualquier serie tiempo de la ciudad del país o de otros países, porque la metodología sería la misma, solo cambiaría el valor de la serie de tiempo.

Los parámetros del modelo (los cuales son generados por Algoritmos Genéticos refiriendonos al cromosoma que es donde están los valores que deben ser extraídos como a_i , b_i y E no dependen de la serie, el error se va calculando mediante el avance del Modelo ARMA y recordando que al individuo que es el mejor de una generación se le van aplicando las operaciones de recombinación y mutación.

Referencias

[Brockwell 1987]

J. Peter Brockwell, Times Series: Theory and Methods, Colorado: Springer, 1987.

[Wellin 2005]

Paul Wellin, An Introduction to Programming with Mathematica, New York: Cambridge, 2005.

[Wellin 2004]

Randy L. Haupt, Practical Genetic Algorithms, New Jersey:Hoboken, 2004.

[Ruskeepää 1987]

H. Ruskeepää, Mathematica Navigator, Filandia: Elsevier, 1987.

[Darwin 1859]

C. Darwin, Origin of species, Inglaterra: John Murray, 1859.

[Jacob 2001]

C. Jacob, Illustrating Evolutionary Computation with Mathematica, Canada: Morgan Kaufman Publishers Murray, 2001.

[WIKIPEDIA 2010]

WIKIPEDIA. Página Secundaria. Estados Unidos. 25 de septiembre del 2010.
http://en.wikipedia.org/wiki/Genetic_algorithm

[STASTIK 2010]

STASTIK. Página Secundaria. Estados Unidos. 25 de septiembre del 2010.
<http://statistik.mathematik.uni-wuerzburg.de/timeseries/download/versions/2006-September-01-times.pdf>

ANEXO

Valores del capítulo 3

Los valores que se muestran a continuación son de la serie de tiempo de la temperatura promedio de la ciudad de Morelia del 1 enero del 2000 hasta el 15 de octubre del 2010.

```
In[102]= WMor = WeatherData["MMMM", "MeanTemperature",
      {{2000, 1, 1}, {2010, 10, 15}}, "Month"]

Out[102]= {{{2000, 1, 4}, 17.16}, {{2000, 2, 3}, 19.08},
  {{2000, 3, 1}, 20.97}, {{2000, 4, 1}, 23.59},
  {{2000, 5, 2}, 23.14}, {{2000, 6, 1}, 21.55},
  {{2000, 7, 1}, 21.7}, {{2000, 8, 2}, 20.91},
  {{2000, 9, 1}, 20.73}, {{2000, 10, 2}, 20.24},
  {{2000, 11, 1}, 19.35}, {{2000, 12, 2}, 17.32},
  {{2001, 1, 2}, 17.48}, {{2001, 2, 1}, 19.43},
  {{2001, 3, 1}, 20.32}, {{2001, 4, 1}, 23.97},
  {{2001, 5, 3}, 22.87}, {{2001, 6, 6}, 20.21},
  {{2001, 7, 1}, 20.25}, {{2001, 8, 1}, 20.92},
  {{2001, 9, 2}, 20.03}, {{2001, 10, 1}, 19.65},
  {{2001, 11, 5}, 17.38}, {{2001, 12, 1}, 17.68},
  {{2002, 1, 2}, 18.57}, {{2002, 2, 1}, 19.08},
  {{2002, 3, 1}, 22.92}, {{2002, 4, 2}, 24.35},
  {{2002, 5, 1}, 24.9}, {{2002, 6, 3}, 22.49},
  {{2002, 7, 1}, 21.29}, {{2002, 8, 1}, 21.19},
  {{2002, 9, 1}, 20.76}, {{2002, 10, 2}, 20.68},
  {{2002, 11, 4}, 16.52}, {{2002, 12, 2}, 17.15},
  {{2003, 1, 2}, 16.88}, {{2003, 2, 7}, 18.26},
  {{2003, 3, 3}, 20.77}, {{2003, 4, 3}, 25.16},
  {{2003, 5, 1}, 26.18}, {{2003, 6, 1}, 22.46},
  {{2003, 7, 1}, 21.03}, {{2003, 8, 3}, 20.96},
  {{2003, 9, 1}, 21.05}, {{2003, 10, 1}, 19.6},
  {{2003, 11, 5}, 20.05}, {{2003, 12, 1}, 17.29},
  {{2004, 1, 4}, 16.82}, {{2004, 2, 2}, 19.22},
  {{2004, 3, 2}, 22.74}, {{2004, 4, 3}, 23.15},
```

{{2004, 5, 2}, 21.13}, {{2004, 6, 1}, 20.52},
{{2004, 7, 4}, 20.71}, {{2004, 8, 1}, 20.88},
{{2004, 9, 3}, 19.86}, {{2004, 10, 1}, 20.56},
{{2004, 11, 1}, 19.59}, {{2004, 12, 6}, 16.15},
{{2005, 1, 4}, 17.49}, {{2005, 2, 6}, 18.23},
{{2005, 3, 1}, 20.23}, {{2005, 4, 1}, 24.45},
{{2005, 5, 1}, 24.12}, {{2005, 6, 1}, 25.28},
{{2005, 7, 1}, 21.53}, {{2005, 8, 1}, 19.6},
{{2005, 9, 3}, 20.69}, {{2005, 10, 2}, 20.14},
{{2005, 11, 1}, 17.55}, {{2005, 12, 2}, 16.83},
{{2006, 1, 3}, 15.92}, {{2006, 2, 3}, 18.96},
{{2006, 3, 1}, 21.65}, {{2006, 4, 2}, 24.48},
{{2006, 5, 1}, 22.33}, {{2006, 6, 1}, 21.94},
{{2006, 7, 2}, 21.03}, {{2006, 8, 1}, 20.09},
{{2006, 9, 6}, 20.59}, {{2006, 10, 2}, 18.69},
{{2006, 11, 1}, 17.64}, {{2006, 12, 1}, 16.5},
{{2007, 1, 1}, 16.91}, {{2007, 2, 1}, 18.38},
{{2007, 3, 1}, 21.28}, {{2007, 4, 3}, 23.42},
{{2007, 5, 1}, 24.04}, {{2007, 6, 3}, 21.69},
{{2007, 7, 1}, 21.38}, {{2007, 8, 2}, 20.25},
{{2007, 9, 9}, 17.91}, {{2007, 10, 1}, 19.81},
{{2007, 11, 4}, 18.25}, {{2007, 12, 17}, 15.5},
{{2008, 1, 12}, 17.41}, {{2008, 2, 1}, 19.44},
{{2008, 3, 1}, 21.25}, {{2008, 4, 1}, 23.92},
{{2008, 5, 1}, 24.31}, {{2008, 6, 2}, 22.44},
{{2008, 7, 1}, 20.4}, {{2008, 8, 1}, 20.05},
{{2008, 9, 1}, 19.59}, {{2008, 10, 1}, 18.48},
{{2008, 11, 1}, 16.15}, {{2008, 12, 1}, 15.62},
{{2009, 1, 1}, 15.99}, {{2009, 2, 1}, 18.31},
{{2009, 3, 1}, 21.09}, {{2009, 4, 1}, 22.96},
{{2009, 5, 1}, 22.48}, {{2009, 6, 1}, 21.83},
{{2009, 7, 1}, 20.81}, {{2009, 8, 1}, 20.25},
{{2009, 9, 1}, 19.32}, {{2009, 10, 1}, 19.08},
{{2009, 11, 1}, 16.3}, {{2009, 12, 1}, 15.51},
{{2010, 1, 1}, 13.82}, {{2010, 2, 1}, 14.7},
{{2010, 3, 1}, 19.24}, {{2010, 4, 1}, 24.25},
{{2010, 5, 1}, 25.14}, {{2010, 6, 1}, 23.24},
{{2010, 7, 1}, 20.56}, {{2010, 8, 1}, 20.4},
{{2010, 9, 1}, 19.83}, {{2010, 10, 1}, 17.65}}