

**ESCUELA SUPERIOR POLITÉCNICA DE  
CHIMBORAZO**



**FACULTAD DE CIENCIAS**

ESCUELA DE FÍSICA Y MATEMÁTICA

TESIS DE GRADO

PREVIA LA OBTENCIÓN

DEL TÍTULO DE INGENIERA EN ESTADÍSTICA

INFORMÁTICA

**“MODELACIÓN Y PRONÓSTICO DEL POTENCIAL  
ENERGÉTICO DEL RIO BLANCO USANDO LA TEORÍA  
DEL CAOS Y UN MÉTODO CONVENCIONAL.”**

PRESENTADO POR:

**AMALIA ISABEL ESCUDERO VILLA**

RIOBAMBA – ECUADOR

2007

**NOMBRE**

**FIRMA**

**FECHA**

**DECANO**

**Dr. Edmundo Caluña**

.....

.....

**DIRECTOR DE ESCUELA**

**Dr. Richard Pachacama**

.....

.....

**DIRECTOR DE TESIS**

**Dr. Arquímedes Haro**

.....

.....

**MIEMBRO DE TRIBUNAL**

**Dr. Msc. Jaime Bejar S.**

.....

.....

**MIEMBRO DE TRIBUNAL**

**Dr. Jorge Congacha**

.....

.....

**DIRECTOR DEL CENTRO DE DOCUMENTACIÓN**

**Sr. Carlos Rodríguez**

.....

.....

**NOTA**

.....

Yo, AMALIA ISABEL ESCUDERO  
VILLA, soy responsable de las  
ideas, doctrinas y resultados  
expuestos en esta Tesis, y el  
patrimonio intelectual de la  
tesis de grado pertenece a la  
ESCUELA SUPERIOR POLITÉCNICA DE  
CHIMBORAZO.

## **AGRADECIMIENTO**

A Dios, por regalarme salud y vida dándome así la oportunidad de superarme y salir siempre adelante.

Mis más sinceros agradecimientos a la EMPRESA ELECTRICA S.A., a mi director de Tesis Dr. Arquímedes Haro, colaboradores Dr. Jaime Bejar, Dr. Jorge Congacha, a mis maestros quienes con dedicación transmitieron sus conocimientos y experiencias durante toda mi carrera y de manera especial al Dr. Celso Recalde director del proyecto BALANCE ENERGÉTICO, DIAGNÓSTICO DEL POTENCIAL DE ENERGÍAS ALTERNATIVAS Y AHORRO ENERGÉTICO DEL CANTÓN RIOBAMBA por el apoyo y financiamiento en la elaboración de éste trabajo.

Agradezco especialmente a mi madre María del Carmen, por el esfuerzo que hizo durante su vida para que yo me convirtiera en una mujer de bien, a mis hermanos Edith, Elsy, Fernando y Cristián por su cariño y apoyo incondicional.

A mis amigos y a todos quienes me apoyaron para que siguiera adelante en mi carrera.

## DEDICATORIA

El presente trabajo va dedicado a Dios, porque sin su voluntad no hubiera sido posible la culminación de la presente tesis.

A mi madre María del Carmen, por todo su amor, comprensión y paciencia.

A mi pequeñita Brianna Lisseth y a mi esposo que siempre estuvieron ahí para darme una palabra de aliento.

*A. Isabel*

# ÍNDICE GENERAL

<i>PORTADA</i> .....	<i>I</i>
<i>AGRADECIMIENTO</i> .....	<i>IV</i>
<i>DEDICATORIA</i> .....	<i>V</i>
<i>ÍNDICE GENERAL</i> .....	<i>VI</i>
<i>ÍNDICE DE ABREVIATURAS</i> .....	<i>X</i>
<i>ÍNDICE DE TABLAS</i> .....	<i>XI</i>
<i>ÍNDICE DE GRÁFICOS</i> .....	<i>XII</i>
<i>ÍNDICE DE ANEXOS</i> .....	<i>XIV</i>
<i>INTRODUCCIÓN</i> .....	<i>XV</i>
<i>OBJETIVOS</i> .....	<i>XVII</i>

## *CAPÍTULO I*

<i>1. FUNDAMENTO TEÓRICO</i> .....	<i>19</i>
<i>1.1. DESCRIPCIÓN DEL SITIO DE ESTUDIO</i> .....	<i>19</i>
<i>1.2. CARACTERÍSTICAS CLIMATOLÓGICAS</i> .....	<i>20</i>
<i>1.2.1. LLUVIAS</i> .....	<i>21</i>
<i>1.2.2. CAUDALES</i> .....	<i>21</i>
<i>1.2.3. VEGETACIÓN</i> .....	<i>22</i>
<i>1.2.4. ACCESOS</i> .....	<i>23</i>
<i>1.3. ENERGÍA</i> .....	<i>23</i>
<i>1.3.1. POTENCIAL ENERGÉTICO</i> .....	<i>24</i>
<i>1.3.2. FUENTES DE ENERGÍA</i> .....	<i>24</i>
<i>1.3.3. DESARROLLO DE LA ENERGÍA HIDROELÉCTRICA</i> .....	<i>24</i>
<i>1.3.4. MICROCENTRALES</i> .....	<i>25</i>
<i>1.3.5. ENERGÍA HIDRÁULICA</i> .....	<i>26</i>
<i>1.3.6. POTENCIAL HIDROELÉCTRICO</i> .....	<i>26</i>
<i>1.4. MODELOS MATEMÁTICOS</i> .....	<i>28</i>
<i>1.5. SERIES TEMPORALES</i> .....	<i>30</i>
<i>1.6. PRONÓSTICOS</i> .....	<i>36</i>
<i>1.7. CLASIFICACIÓN DE LOS MODELOS DE PRONÓSTICO</i> .....	<i>37</i>

1.7.1.	MÉTODOS CUALITATIVOS	37
1.7.2.	MÉTODOS CUANTITATIVOS	38
1.8.	SELECCIÓN DEL MÉTODO DE PRONÓSTICO	40
1.9.	EFICIENCIA EN LOS PRONÓSTICOS	40

## CAPÍTULO II

2.	MÉTODOS DE ANÁLISIS	44
2.1.	MODELOS ARIMA DE Box-Jenkins	44
2.1.1.	GRÁFICA DE LA SERIE	45
2.1.2.	TRANSFORMACIONES	45
2.1.3.	IDENTIFICACIÓN DEL MODELO	45
2.1.3.1.	Autorregresivo $AR(p)$	46
2.1.3.2.	Modelo de media móvil $MA(q)$	47
2.1.3.3.	Modelo Autoregresivo de media móvil $ARMA(p,q)$	47
2.1.3.4.	Modelo Autoregresivo integrado de media móvil $ARIMA(p,d,q)$	48
2.1.4.	SOBREAJUSTE Y SELECCIÓN DEL MODELO	50
2.1.5.	VERIFICACIÓN DEL DIAGNÓSTICO	51
2.1.6.	PREDICCIÓN	51
2.2.	TEORÍA DEL CAOS	51
2.2.1.	CARACTERIZACIÓN DEL CAOS	51
2.2.2.	SISTEMAS CAÓTICOS	54
2.2.3.	CARACTERÍSTICAS GENERALES DE LOS SISTEMAS CAÓTICOS	54
2.2.4.	GEOMETRÍA FRACTAL	55
2.2.5.	ATRACTOR	56
2.2.6.	DIMENSION FRACTAL	58
2.2.7.	TIEMPO DE RETARDO	60
2.2.8.	DIMENSIÓN DE ENCAJE	60
2.2.9.	EXPONENTES DE LYAPUNOV	61

## CAPÍTULO III

3.	RESULTADOS	64
3.1.	MODELO ARIMA DE BOX-JENKINS	64
3.1.1.	GRÁFICA DE LA SERIE	66

3.1.2.	TRANSFORMACIONES	66
3.1.3.	IDENTIFICACIÓN DEL MODELO	71
3.1.4.	SOBREAJUSTE Y SELECCIÓN DEL MODELO	73
3.1.5.	VERIFICACIÓN DEL DIAGNÓSTICO	77
3.1.6.	PREDICCIÓN	81
3.2.	MODELO TISEAN 2.1	83
3.2.1.	TIEMPO DE RETARDO SIN REDUCCIÓN DEL RUIDO	84
3.2.2.	ATRACTOR EXTRAÑO	85
3.2.3.	DIMENSIÓN DE ENCAJE (EMBENDING)	85
3.2.4.	DATOS CON REDUCCIÓN DE RUIDO	86
3.2.5.	TIEMPO DE RETARDO DE SERIE DE DATOS CON REDUCCIÓN DEL RUIDO	87
3.2.6.	ATRACTOR EXTRAÑO CON REDUCCIÓN DEL RUIDO	87
3.2.7.	DIMENSION DE ENCAJE CON REDUCCIÓN DEL RUIDO	88
3.2.8.	ESPECTRO DE LYAPUNOV	89
3.2.9.	PREDICCIÓN	90
CAPÍTULO IV		
4.	ANÁLISIS DE PRECISIÓN	93
4.1.	COMPARACIÓN DE LOS PRONÓSTICOS OBTENIDOS CON EL MODELOS ARIMA Y LA TEORÍA DEL CAOS DURANTE LOS PRIMEROS 14 DÍAS	94
4.1.1.	GRÁFICO DE LAS SERIES	94
4.1.2.	ANÁLISIS DE CORRELACIÓN	95
4.1.3.	ANALISIS DE LOS RESIDUOS O ERRORES	95
4.1.4.	COMPARACIÓN DE MEDIAS DE LAS PREDICCIONES	96
4.2.	COMPARACIÓN DE LOS PRONÓSTICOS OBTENIDOS CON EL MODELOS ARIMA Y LA TEORÍA DEL CAOS DURANTE LOS PRIMEROS 11 DÍAS	97
4.2.1.	GRÁFICO DE LAS SERIES	97
4.2.2.	ANÁLISIS DE CORRELACIÓN	98
4.2.3.	ANALISIS DE LOS RESIDUOS O ERRORES	98
4.3.	PRONÓSTICO DEL POTENCIAL ENERGÉTICO HIDRICO DE RÍO BLANCO	100

<i>CONCLUSIONES</i> -----	102
<i>RECOMENDACIONES</i> -----	104
<i>RESUMEN</i> -----	105
<i>SUMMARY</i> -----	106
<i>BIBLIOGRAFÍA</i> -----	107
<i>BIBLIOGRAFÍA EN INTERNET</i> -----	109
<i>ANEXOS</i> -----	91

## ÍNDICE DE ABREVIATURAS

<b>ACF</b>	Función de Autocorrelación
<b>AR</b>	Auto Regresión
<b>ARIMA</b>	Análisis de Autoregresión Integral y Media Móvil
<b>EERSA</b>	Empresa Eléctrica Riobamba Sociedad Anónima
<b>Lt/s</b>	Litros sobre segundos
<b>m<sup>3</sup>/s</b>	Metros cúbicos sobre segundos
<b>MA</b>	Media Móvil
<b>PACF</b>	Función de Autocorrelación Parcial
<b>X</b>	Variable caudal medio diario

## ÍNDICE DE TABLAS

<i>Tabla 1.- Patrones Teóricos de ACF Y PACF.....</i>	<i>50</i>
<i>Tabla 2.- Estadísticas descriptivas de los caudales medios diarios de río Blanco (2000-2006).....</i>	<i>65</i>
<i>Tabla 3.- Prueba de Levene para la homogeneidad de varianza.....</i>	<i>67</i>
<i>Tabla 4.- Estadísticos del modelo ARIMA(1,1,1)(1,0,1)<sub>12</sub> .....</i>	<i>74</i>
<i>Tabla 5.- Datos Modelo 4 ARIMA (2,1,1)(1,0,1)<sub>12</sub> .....</i>	<i>75</i>
<i>Tabla 6.- Prueba de Levene de homogeneidad de varianza de los residuos o errores ARIMA (1,1,1)(1,0,1)<sub>12</sub> y ARIMA (2,1,1)(1,0,1)<sub>12</sub> .....</i>	<i>76</i>
<i>Tabla 7.- Ajuste del modelo ARIMA (1,1,1)(1,0,1)<sub>12</sub>.....</i>	<i>77</i>
<i>Tabla 8.- Estadísticas Descriptivas de los Residuos.....</i>	<i>78</i>
<i>Tabla 9.- Prueba cero para la comparación de la media de los residuos.....</i>	<i>79</i>
<i>Tabla 10.- Prueba de homogeneidad de varianza .....</i>	<i>79</i>
<i>Tabla 11.- Prueba de Kolmogorov-Smirnov para una muestra .....</i>	<i>81</i>
<i>Tabla 12.- Correlación de los datos reales y predichos durante 17 días.....</i>	<i>82</i>
<i>Tabla 13.- Correlación de los datos reales y predichos. ....</i>	<i>90</i>
<i>Tabla 14.- Prueba de Tukey para comparación de medias durante 17 días....</i>	<i>93</i>
<i>Tabla 15.- Correlación de los datos reales y predichos. ....</i>	<i>95</i>
<i>Tabla 16.- Prueba de Tukey para comparación de medias durante 14 días de pronóstico.....</i>	<i>96</i>
<i>Tabla 17.- Correlación de los datos reales y predichos. ....</i>	<i>98</i>
<i>Tabla 18.- Prueba de Tukey para comparación de medias durante 11 días de pronóstico.....</i>	<i>99</i>
<i>Tabla 19.- Prueba de comparación de medias de pronósticos a un período de 12 días. ....</i>	<i>99</i>
<i>Tabla 20.- Predicciones con el modelo ARIMA y teoría del CAOS del potencial energético hídrico.....</i>	<i>100</i>
<i>Tabla 21.- Prueba de Tukey para comparación de las medias del pronóstico del potencial energético hídrico del río Blanco .....</i>	<i>101</i>

## ÍNDICE DE GRÁFICOS

<i>GRAFICO 1.- Simatic S5 tablero de control y procesador para la medición los caudales río Blanco.....</i>	<i>22</i>
<i>GRAFICO 2.- Vegetación de los alrededores de la central hidroeléctrica de río Blanco.....</i>	<i>23</i>
<i>GRAFICO 3.- Presa de Itaipú, Paraguay.....</i>	<i>25</i>
<i>GRAFICO 4.- Caída de agua río Blanco.....</i>	<i>26</i>
<i>GRAFICO 5.- Representación gráfica de una Serie Temporal.....</i>	<i>30</i>
<i>GRAFICO 6.- Triángulo de Sierpinski. Hoja de Helecho.....</i>	<i>55</i>
<i>GRAFICO 7.- Representación visual de atractor.....</i>	<i>57</i>
<i>GRAFICO 8.- Serie temporal de caudales de río Blanco. Período 2000 - 200666</i>	
<i>GRAFICO 9.- Diagrama de Box Plot.....</i>	<i>67</i>
<i>GRAFICO 10.- Visualización de los caudales medios diarios y su media.....</i>	<i>68</i>
<i>GRAFICO 11.- Correlograma simple de la serie con los 640 primeros retardos. ....</i>	<i>69</i>
<i>GRAFICO 12.- Autocorrelación simple de los primeros 40 rezagos o retardos</i>	<i>70</i>
<i>GRAFICO 13.- Función de Autocorrelación simple sobre los primeros 40 retardos de río Blanco 2000-2006.....</i>	<i>71</i>
<i>GRAFICO 14.- Función de autocorrelación parcial sobre los primeros 40 retardos de río Blanco 2000-2006.....</i>	<i>72</i>
<i>GRAFICO 15.- Errores que se obtienen como diferencia entre los valores de la serie de caudales originales y la resultante del modelo.....</i>	<i>77</i>
<i>GRAFICO 16.- Diagrama de autocorrelación simple de los errores o residuos</i>	<i>80</i>
<i>GRAFICO 17.- Diagrama de autocorrelación parcial de los errores o residuos</i>	<i>80</i>
<i>GRAFICO 18.- Datos de los caudales originales y la predicción ARIMA durante 17 días. ....</i>	<i>82</i>
<i>GRAFICO 19.- Promedios diarios de los caudales de río Blanco (2000-2007).</i>	<i>83</i>
<i>GRAFICO 20.- Tiempo de retardo de los datos de caudales sin reducción del ruido.....</i>	<i>84</i>

<i>GRAFICO 21.- Proyección bidimensional del atractor extraño de los datos de caudal sin reducción del ruido.....</i>	<i>85</i>
<i>GRAFICO 22.- Gráfico de la dimensión de encaje de los datos de caudal sin reducción del ruido.....</i>	<i>85</i>
<i>GRAFICO 23.- Gráfico de los datos de caudal con reducción del ruido. ....</i>	<i>86</i>
<i>GRAFICO 24: Gráfico de la función de retardo (mutual) que determina el tiempo de retardo para la reconstrucción del espacio de las fases de los datos de caudal con reducción del ruido.....</i>	<i>87</i>
<i>GRAFICO 25.- Proyección bidimensional del atractor extraño de los datos de caudales con reducción del ruido.....</i>	<i>87</i>
<i>GRAFICO 26.- Gráfico de la función de encaje (falsos vecinos) que determina la dimensión de encaje para la reconstrucción del espacio de las fases de los datos de caudales con reducción del ruido. ....</i>	<i>88</i>
<i>GRAFICO 27.- Gráfico del espectro de Lyapunov de los datos de caudales con reducción del ruido.....</i>	<i>89</i>
<i>GRAFICO 28.- Datos reales y predichos de los caudales de río Blanco durante los 14 días.....</i>	<i>90</i>
<i>GRAFICO 29.- Gráfico de los datos reales &amp; predicción ARIMA &amp; Teoría del Caos en los primeros 14 días.....</i>	<i>94</i>
<i>GRAFICO 30.- Gráfico de los datos reales &amp; predicción ARIMA &amp; Teoría del Caos en los primeros 11 días.....</i>	<i>97</i>
<i>GRAFICO 31.- Pronósticos de potencial energético hídrico del río Blanco usando la TEORÍA DEL CAOS y el modelo ARIMA.....</i>	<i>101</i>

## ÍNDICE DE ANEXOS

<i>ANEXO 1: Caudales reales y predichos, modelo ARIMA y Teoría del CAOS durante 17 días .....</i>	<i>110</i>
<i>ANEXO 2: Central Hidroeléctrica río Blanco Quimiag – Riobamba.....</i>	<i>110</i>

## INTRODUCCIÓN

Este trabajo de tesis se desarrolló con el afán de predecir el potencial energético hídrico de río Blanco de Quimiag cuyos resultados obtenidos en esta investigación son parte importante para el proyecto en vigencia de “Balance Energético y Diagnóstico del Potencial de Energías Alternativas del Cantón Riobamba” financiado por el CONESUP, que tiende a resolver problemas graves relacionados con el mal uso y aprovechamiento de la energía, y con más razón cuando se conoce que en el centro del país tiene el costo más elevado por kW hora, así como deficiencia en planificación y no existen políticas que fomenten las energías renovables, como la hídrica; además, no existe una base de datos históricos adecuadamente procesados usando los conocimientos actuales que se han desarrollado en el campo de las matemáticas, no se tienen análisis de factibilidad, los inversionistas tienen desconfianza, a esto se suma la pobreza y pocas perspectivas de desarrollo.

La capacidad de predecir con certeza y precisión el comportamiento futuro de un sistema, sea físico, biológico o social se ha situado como una de las máximas aspiraciones de la ciencia clásica y al conocer sobre el desarrollo de otros métodos de predicción como la teoría del caos que buscan mayor precisión surge la necesidad de compararlo con modelos de trayectoria relevante como son los modelos ARIMA de BOX-Jenkins (Autorregresive Integrated Moving Average).

Para el análisis del modelo ARIMA se utilizó el software SPSS 11 y TISEAN 2.1 para el análisis con la teoría del caos con el fin de identificar las características

necesarias para el ajuste del modelo a los datos originales esperando pronósticos confiables. La precisión de pronóstico se halla mediante el análisis de los residuos o errores, correlación, prueba de Tukey de comparación de medias. El período de predicción con correlación significativa es de 14 días con la teoría del caos y 17 días con el modelo ARIMA.

A pesar que los residuos o errores obtenidos en las predicciones con la teoría del caos son menores que los obtenidos con el modelo ARIMA la prueba de Tukey muestra que no existe diferencia significativa entre las predicciones y los datos reales hasta un período de pronóstico de 11 días (tabla 18); es decir, las medias de los pronósticos del potencial energético hídrico de Río Blanco hallados con los caudales predichos con los dos métodos son estadísticamente iguales al potencial hallado con los datos reales durante un periodo de 11 días al 95% de confiabilidad, a partir del doceavo día existe diferencia significativa en las predicciones; es decir, la teoría del Caos es más precisa que el modelo ARIMA a partir del doceavo día de pronóstico (tabla 19).

En nuestro país, el uso de este tipo de herramienta matemática es incipiente, debido al desarrollo mismo del país y particularmente de las ciencias exactas, a pesar que la misma podría ser una solución importante para mucho de los problemas que nos aquejan, por ello la Empresa Eléctrica Riobamba S.A debería implementar una base de datos informática para un mejor manejo de los datos, y de esta forma fomentar estudios de potencial hidroeléctrico mediante modelos de pronósticos para su mejor aprovechamiento.

# OBJETIVOS

## GENERAL

Modelar y predecir el potencial energético hídrico que posee el río Blanco del cantón Riobamba.

## ESPECÍFICOS

- Procesar los datos hidrográficos para determinar su potencial energético.
- Seleccionar el método convencional para el pronóstico.
- Pronosticar el potencial hídrico usando un método convencional.
- Determinar el tiempo de retardo de los datos de potencial hídrico.
- Determinar la dimensión de encaje de los datos del potencial hídrico.
- Determinar los coeficientes de Lyapunov de los datos del potencial hídrico.
- Pronosticar el potencial hídrico usando la teoría de Caos
- Analizar la precisión de los dos métodos.

**CAPÍTULO I**

**FUNDAMENTO**

**TEÓRICO**

# 1. FUNDAMENTO TEÓRICO

## 1.1. DESCRIPCIÓN DEL SITIO DE ESTUDIO

El río Blanco está en la Provincia de Chimborazo, cantón Riobamba en la parroquia de Quimiag; forma parte del sistema hidrográfico del río Pastaza, corre de sur-orienta a nor-occidente desembocando en el chambo, en el sector de puntal y tiene su nacimiento en las estribaciones del volcán – nevado El Altar.

Las cuencas del río Blanco, es de forma trapezoidal, está ubicado entre  $78^{\circ} 26'$  y  $78^{\circ} 31,5'$  de longitud occidental y  $1^{\circ} 40'$  y  $1^{\circ} 46'$  de latitud sur y tiene una superficie aproximada de  $63 \text{ km}^2$  hasta el sitio de confluencia con una quebrada sin nombre que entra por el margen izquierdo.

Orográficamente, la cuenca tiene un aspecto muy accidentado; su borde oriental está aproximadamente a 4 700 m de altitud, llegando la esquina nor-oriental a tocar la zona nevada que rodea casi totalmente al cráter del volcán, y el sitio de toma está en la cota 3 400 por no disponer de mapa de curvas de nivel que cubran la cuenca es imposible determinar su elevación media, a partir de la curva hipsométrica; tampoco se puede calcular el rectángulo equivalente, ni la pendiente de drenaje según el método de Horton. Sin embargo por diferencia de cotas se puede obtener la pendiente media del cause principal, partiendo desde los orígenes de la quebrada Pailacajas cuya altitud

aproximada es de 4 650 m el valor así obtenido para el cause de 10,5 km de largo lo que confirma que el río y sus quebradas afluentes corren con régimen torrencial.

## **1.2. CARACTERÍSTICAS CLIMATOLÓGICAS**

La precipitación media anual es de 1 073,4 mm en la cuenca del río Blanco, con valores de 668,4 mm y 1 879,1 mm para los años secos y húmedos respectivamente dentro del mismo período. La distribución de la lluvia durante el año medio se produce de modo que se tiene el valor más bajo en el mes de agosto, el más alto entre marzo y abril y otros dos valores altos en junio y octubre.

Los otros factores climáticos de la cuenca pueden asumirse en base a los datos de estaciones ubicadas en sitios semejantes, sobre todo en lo que se refiere a la altitud. La temperatura media anual debe aproximarse a los 6 °C y la evaporación anual a los 250 mm como promedio, ambas influenciadas por la cercanía de la nieve. El clima de la cuenca puede calificarse como semihúmedo de páramo.

### **1.2.1. LLUVIAS**

Los vientos húmedos provenientes de la hoya amazónica, que descargan su humedad en las partes altas y que penetran hacia el callejón interandino siguiendo el cañón de Pastaza y otros cauces, ejercen una influencia marcada en el régimen de lluvias de la zona, cuyo análisis y estudio conduce a determinar los volúmenes de agua precipitada sobre la cuenca del Blanco, mediante el método de las isoyetas, ya que no se tiene datos hidrométricos.

La ubicación del país en la zona ecuatorial, implica contar con sectores de baja presión y con cambios de los centros de presión atmosférica, lo que produce una gran variación interanual de las lluvias.

### **1.2.2. CAUDALES**

El caudal es el volumen de agua que fluye a través de una sección de un curso de agua por unidad de tiempo, Se mide en metros cúbicos por segundo ( $m^3/s$ ).

Simatic S5 es el procesador para la medición de caudales en (L/s) de la central en río Blanco.



**GRAFICO 1.-** Simatic S5 tablero de control y procesador para la medición los caudales río Blanco.

### **1.2.3. VEGETACIÓN**

La cobertura vegetal influye en el escurrimiento de las aguas lluvias. La mayor parte de la cuenca tiene pajonales de páramo; la zona mas alta está desprovista de toda vegetación por las pendientes fuertes, por el suelo rocoso y por la cercanía a la nieve; tan solo junto a los causes y cerca del sitio de toma hay pequeñas manchas de vegetación arbustiva.

En las partes bajas y relativamente planas hay sembríos, mientras que en las zonas con pendientes fuertes y zonas de páramo se encuentra una vegetación natural, representada por árboles, chaparros y pajonales.



**GRAFICO 2.-** Vegetación de los alrededores de la central hidroeléctrica de río Blanco.

#### **1.2.4. ACCESOS**

A la zona se puede llegar por vía carrozable desde Riobamba asta la Central Hidroeléctrica de río Blanco en Quimiag.

### **1.3. ENERGÍA**

La energía se define en forma general como toda causa capaz de producir un trabajo, y su manifestación es precisamente la realización de su virtualidad, es decir, la producción de un trabajo, o bien su transformación en otra forma de energía.

### **1.3.1. POTENCIAL ENERGÉTICO**

Es la cantidad potencialmente aprovechable de energía presente en la naturaleza, independiente de cuál sea la fuente energética, posible de ser aprovechada mediante el uso de tecnología.

### **1.3.2. FUENTES DE ENERGÍA**

La energía se produce en diferentes fuentes y es almacenada en distintas formas. Las fuente se pueden clasificar en *primarias y secundarias*, según pueda obtenerse de ellas la energía directamente o sea necesario recurrir a otra fuente, así por ejemplo la energía eléctrica es una fuente secundaria de energía porque para su producción es necesario recurrir a otra fuente de energía. Las fuentes de energía se clasifican en *renovables y no renovables*, según que su energía se siga produciendo en la actualidad y su consumo sea repuesto. O que ya no se produzca y su consumo acabe por agotar la reserva. A éste segundo tipo pertenecen los fósiles: petróleo, carbón y gas natural; en cambio son renovables la energía solar, eólica, **hidráulica**, biomasa y la debida a mareas, olas y gradientes térmicos permanentes.

### **1.3.3. DESARROLLO DE LA ENERGÍA HIDROELÉCTRICA**

La primera central hidroeléctrica se construyó en 1880 en Northumberland, Gran Bretaña. El renacimiento de la energía hidráulica se produjo por el

desarrollo del generador eléctrico, seguido del perfeccionamiento de la turbina hidráulica y debido al aumento de la demanda de electricidad a principios del siglo XX. En 1920 las centrales hidroeléctricas generaban ya una parte importante de la producción total de electricidad.



**GRAFICO 3.-** Presa de Itaipú, Paraguay

#### **1.3.4. MICROCENTRALES**

Las microcentrales hidroeléctricas aprovechan los recursos hídricos sin deteriorar el medio ambiente. Sus obras civiles son de bajo impacto ambiental pues en la mayoría de casos no requieren de represas que inundan tierras fértiles o de reserva natural. Por el contrario promueven la conservación de las cuencas ya que crean conciencia en los usuarios de la importancia del agua y su conservación debido a la íntima relación que existe entre el caudal de agua y la cantidad de energía generada. La microhidroenergía hace uso de un recurso natural renovable de una forma sostenible.

### **1.3.5. ENERGÍA HIDRÁULICA**

Energía que se obtiene de la caída del agua desde cierta altura hasta un nivel inferior, lo que provoca el movimiento de ruedas hidráulicas o turbinas. La hidroelectricidad es un recurso natural disponible en las zonas que presentan suficiente cantidad de agua. Su desarrollo requiere construir pantanos, presas, canales de derivación, y la instalación de grandes turbinas y equipamiento para generar electricidad.



**GRAFICO 4.-** Caída de agua río Blanco

### **1.3.6. POTENCIAL HIDROELÉCTRICO**

Los recursos hidroeléctricos son una fuente de energía permanentemente renovable. Su valoración cuantitativa se realiza con el auxilio del concepto de potencial hidroeléctrico, es decir, de la energía de un curso de agua. Se distinguen tres formas de potencial hidroeléctrico:

- **Teórico (bruto):** energía teórica del curso del agua (sin considerar las pérdidas).
- **Técnico:** energía del curso del agua, que técnicamente puede utilizarse tomando en consideración las pérdidas.
- **Económico:** energía del curso del agua, cuya utilización resulta efectiva económicamente.

**Concepto de potencial hidroeléctrico:** Cantidad total de energía eléctrica de una cuenca hidrográfica, posible de ser aprovechada mediante tecnología. En las grandes cuencas se mide en MW/año (MegaWatts por año).

La potencia obtenida a través de los recursos hidráulicos depende del volumen de agua que fluye por unidad de tiempo y de la altura de caída de ésta. Una central hidroeléctrica es un conjunto de obras destinadas a convertir la energía cinética y potencial del agua, en energía utilizable como es la electricidad. Esta transformación se realiza a través de la acción que el agua ejerce sobre una turbina hidráulica, la que a su vez le entrega movimiento rotatorio a un generador eléctrico.

La potencia eléctrica nominal de una central hidroeléctrica viene modelada en forma aproximada por la ecuación:

$$P = Q_e \cdot H_n \cdot k \quad (1.1)$$

Donde:

$P$  = *Potencia*

$Q_e$  = *Caudal de equipamiento*

$H_n$  = *Salto neto (284,27 m)*

$k$  = *Constante (7,84 kg / m<sup>2</sup> s<sup>2</sup>)*

Pero como el caudal turbinable es una variable aleatoria, y aunque con el pequeño embalse se pueda regular en cierta medida, se seguirá teniendo una variable aleatoria, por lo que se debe evaluar la potencia para cada uno de los caudales medios diarios. Se halló la potencia con cada uno de los caudales medios diarios, con una previa asignación de un caudal de equipamiento, y se calculo la energía para cada uno de los días.

#### **1.4. MODELOS MATEMÁTICOS**

En el mundo real existen problemas (o fenómenos) que pueden ser analizados y resueltos empleando la matemática, una herramienta que interpreta estas situaciones son los Modelos Matemáticos.

La mayoría de problemas que requieren de alguna toma de decisión pueden ser expresados en términos matemáticos, lo cual permitirá analizar los objetivos o metas que se pretenden alcanzar.

El proceso de análisis se realiza utilizando un modelo (o representación limitada de la realidad), donde factores decisivos o variables se combinan de manera lógica, los mismos que permitirán explicar el problema (o fenómeno),

**Modelo matemático.-** Se define como una descripción desde el punto de vista de las matemáticas de un hecho o fenómeno del mundo real, desde el tamaño de la población, hasta fenómenos físicos. El objetivo del modelo matemático es entender ampliamente el fenómeno y tal vez predecir su comportamiento en el futuro.

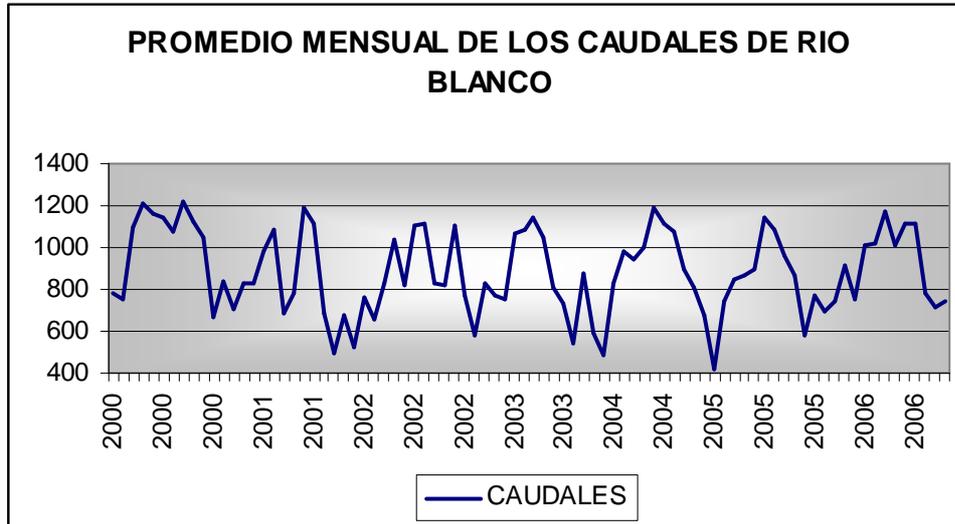
El proceso para elaborar un modelo matemático es el siguiente:

1. Encontrar un problema del mundo real
2. Formular un modelo matemático acerca del problema, identificando variables (dependientes e independientes) y estableciendo hipótesis lo suficientemente simples para tratarse de manera matemática.
3. Aplicar los conocimientos matemáticos que se posee para llegar a conclusiones matemáticas.
4. Comparar los datos obtenidos como predicciones con datos reales. Si los datos son diferentes, se reinicia el proceso.

Es importante mencionar que un modelo matemático no es completamente exacto con problemas de la vida real, de hecho, se trata de una idealización.

## 1.5. SERIES TEMPORALES

Llamamos “Serie de Tiempo” a un conjunto de mediciones de cierto fenómeno o experimento registradas secuencialmente en el tiempo.



Fuente: EERSA

**GRAFICO 5.-** Representación gráfica de una Serie Temporal

Estas observaciones serán denotadas por:

$$\{Y(t_1), Y(t_2), \dots, Y(t_n)\} = \{Y(t) : t \in T \subseteq R\} \quad (1.2)$$

con  $Y(t_i)$  el valor de la variable  $x$  en el instante  $t_i$ . Por ejemplo a cada hora, mensualmente, trimestralmente, semestralmente, etc.

Un modelo clásico de una serie de tiempo, puede ser expresada como suma ( $Y_t = \bar{Y}_t + Y_t^s$ ) o producto ( $Y_t = \bar{Y}_t \cdot Y_t^s$ ) de las componentes: Tendencia, Ciclo, Estacionalidad y Residuo o error.

### ***a) Tendencia***

Movimiento de larga duración que se mantiene durante el período de observación y puede ajustarse a funciones matemáticas (recta, parábola, etc.)

### ***b) Estacionalidad***

Movimiento que se produce en una serie, dentro de un período anual, por motivos climáticos y, con base en éstos, por motivos económicos de organización social (recolecciones, vacaciones, etc.) El período puede ser también de un mes (las ventas de los comercios disminuyen a final de mes y aumentan al principio), la semana (la productividad de los trabajadores suele ser menor en primer y último día de la semana) o incluso el día (cotizaciones en la Bolsa).

### ***c) Ciclo***

Oscilaciones alrededor de la tendencia producidos por períodos alternativos de prosperidad y depresión. Normalmente en una serie económica se superponen distintos ciclos de esta clase, lo que hace que en la práctica ésta sea la componente más difícil de determinar.

**d) Residuo o error:**

Oscilaciones erráticas que aún quedan en la serie original tras eliminar los 3 anteriores componentes

El análisis de las series temporales se enmarca en los procesos estocásticos. Una sucesión de variables aleatorias  $Y_t$  ( $t=1, 2, \dots, n$ ) donde cada valor observado puede considerarse una muestra aleatoria de tamaño uno de la variable. Además  $Y_a$  y  $Y_b$  están separadas por  $k$  retardos  $|a - b| = k$ .

Un **proceso estocástico** es una sucesión de variables aleatorias  $Y_t$  ordenadas, pudiendo tomar  $t$  cualquier valor entre  $-\infty$  y  $\infty$ . Por ejemplo, la siguiente sucesión de variables aleatorias puede ser considerada como proceso estocástico:

$$Y_{-5}, Y_{-4}, Y_{-3}, Y_{-2}, \dots, Y_3, Y_4 \quad (1.3)$$

El subíndice  $t$  no tiene, en principio, ninguna interpretación a priori, aunque si hablamos de proceso estocástico en el contexto del análisis de series temporales este subíndice representará el paso del tiempo.

Cada una de las variables  $Y_t$  que configuran un proceso estocástico tendrá su propia función de distribución con sus correspondientes momentos. Así mismo, cada par de esas variables tendrán su correspondiente función de distribución conjunta y sus funciones de distribución marginales. Esto mismo ocurrirá, ya no

para cada par de variables, sino para conjuntos más amplios de las mismas. De esta forma, para caracterizar un proceso estocástico deberíamos especificar las funciones de distribución conjunta de cualquier conjunto de variables:

$$Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_m}, \quad (1.4)$$

cualesquiera que fueran los valores de  $(t_1, t_2, \dots, t_m)$  y cualquiera que fuera el valor de  $m$ ; por ejemplo:

$$\begin{aligned} Y_1, Y_2, Y_{31}, (t_1 = 1 \text{ y } m = 3) \\ Y_3, Y_4, Y_5, Y_{61}, (t_1 = 3 \text{ y } m = 4) \end{aligned} \quad (1.5)$$

Habitualmente, conocer esas funciones de distribución resulta complejo de forma que, para caracterizar un proceso estocástico, basta con especificar la media y la varianza para cada  $Y_t$  y la covarianza para variables referidas a distintos valores de  $t$ :

$$\begin{aligned} E[Y_t] &= \mu_t \\ \sigma_t^2 &= \text{Var}(y_t) = E[Y_t - \mu_t]^2 \\ \gamma_t &= \text{Cov}(Y_y, Y_5) = E[(Y_t - \mu_t)(Y_5 - \mu_5)] \end{aligned} \quad (1.6)$$

Las distribuciones de probabilidad podrían no estar completamente caracterizadas en algunas de las variables, los momentos podrían no coincidir incluso no existir para alguna de las variables aleatorias, lo mismo puede ocurrir con las distribuciones conjuntas o marginales. Sin embargo, de todos los tipos de procesos estocásticos posibles, nos interesan especialmente dos de ellos a los que la estadística ha dado nombres precisos:

- **Ruido blanco** es una sucesión de variables aleatorias (proceso estocástico) con esperanza (media) cero, varianza constante e independientes para distintos valores de  $t$  (covarianza nula).
- **Proceso estocástico estacionario.** Decimos que un proceso estocástico es estacionario si las funciones de distribución conjuntas son invariantes con respecto a un desplazamiento en el tiempo (variación de  $t$ ). Es decir, considerando que  $t, t+1, t+2, \dots, t+k$  reflejan períodos sucesivos:

$$F(Y_t, Y_{t+1}, \dots, Y_{t+k}) = F(Y_{t+m}, Y_{t+1+m}, \dots, Y_{t+k+m}) \quad (1.7)$$

para cualquier  $t$ ,  $k$  y  $m$ ; por ejemplo:

$$F(Y_1, Y_2, \dots, Y_6) = F(Y_{10}, Y_{11}, \dots, Y_{15}) \quad (1.8)$$

Donde  $t = 1$ ,  $k = 5$ ,  $m = 9$

$$F(Y_3, Y_4, Y_5) = F(Y_7, Y_8, Y_9) \quad (1.9)$$

Donde  $t = 3$ ,  $k = 2$ ,  $m = 4$

$$F(Y_3, Y_4, Y_5) = F(Y_7, Y_8, Y_9) \quad (1.10)$$

Esta definición de estacionariedad se conoce como estacionariedad en sentido estricto o fuerte y puede relajarse sustancialmente utilizando la denominada estacionariedad en sentido amplio o débil. Decimos que un proceso estocástico es débilmente estacionario si:

- Las esperanzas matemáticas de las variables aleatorias no dependen del tiempo, son constantes:

$$E[T_t] = E[Y_{t+m}] \quad \forall m \quad (1.11)$$

- Las varianzas tampoco dependen del tiempo y son finitas:

$$Var[T_t] = Var[Y_{t+m}] \neq \infty \quad \forall m \quad (1.12)$$

- Las covarianzas entre dos variables aleatorias del proceso correspondientes a períodos distintos de tiempo (distintos valores de t) solamente dependen del lapso de tiempo transcurrido entre ellas:

$$Cov[T_t, Y_s] = Cov[Y_{t+m}, Y_{s+m}] \quad \forall m \quad (1.13)$$

De esta última condición se desprende que, si un fenómeno es estacionario, sus variables pueden estar relacionadas linealmente entre si, pero de forma que la relación entre dos variables sólo depende de la distancia temporal k transcurrida entre ellas.

Lógicamente, la estacionariedad en sentido estricto garantiza la estacionariedad en sentido amplio pero no al revés.

Una vez introducido el concepto genérico de proceso estocástico puede decirse que una serie temporal cualquiera es, en realidad, una muestra, una realización concreta con unos valores concretos de un proceso estocástico teórico, real. El análisis de series que vamos a estudiar tratará, a partir de los datos de una serie temporal, inferir las características de la estructura probabilística subyacente, del verdadero proceso estocástico.

## 1.6. PRONÓSTICOS

Pronosticar es emitir un enunciado sobre lo que es probable que ocurra en el futuro, basándose en análisis y en consideraciones de juicio, son premisas o suposiciones básicas en que se basan la planeación y la toma de decisiones.

Muchos métodos de pronósticos se apoyan en técnicas matemáticas complejas; el pronóstico se necesita como elemento de otros modelos y algunos pronósticos son una ayuda esencial en la planeación y solución de problemas.

En realidad, los pronósticos no sólo se utilizan como elemento de los modelos de solución de problemas mediante la ciencia administrativa, sino que establecen además las premisas a partir de las cuales se elaboran los planes y controles.

**Pronóstico cuantitativo.** Este hace una extrapolación del pasado o se utiliza cuando se cuenta con suficientes datos estadísticos o confiables para especificar las relaciones existentes entre variables fundamentales.

El pronóstico basado en la extrapolación, como un análisis de series de tiempo, recurre a las tendencias pasadas o presentes a fin de proyectar los acontecimientos futuros. Así, los registros de ventas en los últimos años podrían servir para proyectar el patrón de ventas para el próximo año.

## 1.7. CLASIFICACIÓN DE LOS MODELOS DE PRONÓSTICO



### 1.7.1. MÉTODOS CUALITATIVOS

Las técnicas cualitativas se usan cuando los datos son escasos, por ejemplo cuando se introduce un producto nuevo al mercado.

Estas técnicas usan el criterio de la persona y ciertas relaciones para transformar información cualitativa en estimados cuantitativos.

a) **Método Delphi.**- Se usa para pronósticos a largo plazo, pronósticos de ventas de productos nuevos y pronósticos tecnológicos.

- Tiempo estimado, más de dos meses.
- Exactitud, de regular a muy buena.

b) **Investigación de Mercados.** Se usa para evaluar y probar hipótesis acerca de mercados reales.

- Tiempo estimado, más de tres meses.

- Exactitud, puede ser excelente, dependiendo del cuidado que se haya puesto en el trabajo.

c) **Consenso de un Panel.** Tiene los mismos usos que el Método Delphi.

- Tiempo estimado, más de dos semanas.
- Exactitud, de baja a regular.

d) **Pronósticos Visionarios.** Se usa para hacer una profecía del futuro usando la intuición personal.

- Tiempo estimado, una semana.
- Exactitud, mala.

e) **Analogía Histórica.** Se usa para productos nuevos, basándose en el análisis comparativo de la introducción y crecimiento de productos similares.

- Tiempo estimado, más de un mes.
- Exactitud, de buena a regular.

### 1.7.2. MÉTODOS CUANTITATIVOS

Análisis de series de tiempo. El análisis consiste en encontrar el patrón del pasado y proyectarlo al futuro.

Patrones de una serie de tiempo:

- Horizontal o estacionario
- Tendencia a largo plazo
- Efecto estacional
- Efecto cíclico

a) **Métodos de proyección.** Estos métodos tratan de encontrar el patrón total de los datos para proyectarlos al futuro, y son:

- Promedios Móviles
- Suavización Exponencial
- Box-Jenkins

b) **Método de separación.** Es aquel que separa la serie en sus componentes para identificar el patrón de cada componente, y se llama, Método de Descomposición de Series de Tiempo.

- Modelos Causales
- Modelos de Regresión
- Regresión lineal simple
- Regresión lineal múltiple

c) **Modelos Econométricos.** Un modelo econométrico es un sistema de ecuaciones de regresión interdependientes que describe algún sector de actividades económicas, ventas o utilidades.

## **1.8. SELECCIÓN DEL MÉTODO DE PRONÓSTICO**

Para realización de un pronóstico se debe tomar en cuenta los siguientes factores:

- a) El contexto del pronóstico
- b) La relevancia y disponibilidad de datos históricos
- c) El grado de exactitud deseado
- d) El periodo de tiempo que se va a pronosticar
- e) El análisis de costo-beneficio del pronóstico
- f) El punto del ciclo de vida en que se encuentra el producto.

## **1.9. EFICIENCIA EN LOS PRONÓSTICOS**

Hay quien pueda afirmar, el modelo es útil solo en la medida que anticipa la evolución de la variable. En este sentido se está aceptando que uno espere a que vengan las futuras observaciones para después analizar la calidad del modelo. Esto se le llama una evaluación ex-post.

Uno puede hacer una evaluación ex-ante, o sea, no incluir las últimas 1, 2, 3,... en total guardar "s" observaciones de la serie mientras construye al modelo, con el modelo calculado en la sub-muestra uno proyecta los valores y los compara con los "s" datos que separó, esta acción permite analizar el error de pronóstico y evaluar al modelo, si los errores son pequeños el modelo es aceptable.

Una vez terminado el análisis entre datos proyectados versus datos reales. Se incluyen estas últimas observaciones y se vuelve a estimar el modelo con la muestra completa y generar sus pronósticos.

El error de pronóstico se define como:

$$e_j = Y_{t+j} - \hat{Y}_t(j) \quad (1.14)$$

Donde:

$Y_t(j)$  es el pronóstico de  $Y_{t+j}$  realizado al momento t para j periodos adelante,

los criterios de pronóstico son los siguientes:

1. Error porcentual medio:

$$EPM = \left( \frac{1}{s} \sum_{j=1}^s \frac{e_j}{Y_{t+j}} \right) * 100\% \quad (1.15)$$

2. Error medio cuadrático:

$$EMC = \frac{\sum_{j=1}^s e_j^2}{s} \quad (1.16)$$

3. Error medio Absoluto:

$$EMA = \frac{\sum_{j=1}^s |e_j|}{s} \quad (1.17)$$

4. Error absoluto porcentual medio:

$$EAPM = \left( \frac{1}{s} \sum_{j=1}^s \left| \frac{e_j}{Y_{t+j}} \right| \right) * 100\% \quad (1.18)$$



# **CAPÍTULO II**

# **METODOLOGÍA**

## **2. MÉTODOS DE ANÁLISIS**

### **2.1. MODELOS ARIMA DE Box-Jenkins**

En 1970, Box y Jenkins desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental.

Una parte importante de esta metodología está pensada para liberar al investigador econométra de la tarea de especificación de los modelos dejando que los propios datos temporales de la variable a estudiar nos indiquen las características de la estructura probabilística subyacente. En parte, los procedimientos que vamos a analizar se contraponen a la "forma tradicional" de identificar y especificar un modelo apoyándonos en las teorías subyacentes al fenómeno analizado aunque, convenientemente utilizados, los conceptos y procedimientos que examinaremos constituyen una herramienta útil para ampliar y complementar los conocimientos econométricos básicos. Cada observación en un momento dado es modelada en función de los valores anteriores.

Dicho análisis consta de las siguientes fases o etapas:

### **2.1.1. GRÁFICA DE LA SERIE**

Con el graficador de SPSS se representa la serie histórica a fin de decidir la estacionariedad de la serie.

### **2.1.2. TRANSFORMACIONES**

En muchos casos conviene, antes de aplicar el método, transformar la serie original  $Y$ , según una serie de alternativas que ofrece el programa. Esta transformación es imprescindible en caso de que la serie no sea estacionaria.

Un proceso estocástico es estacionario si:

- La media de  $Y_t$  es constante
- La varianza de  $Y_t$  es constante
- La correlación entre  $Y_t$  y  $Y_{t+k}$  depende únicamente del número de retardos  $k$ .

### **2.1.3. IDENTIFICACIÓN DEL MODELO**

Los análisis se basan en un modelo explícito. Estos modelos se conocen con el nombre genérico ARIMA (Auto Regresive Integrated Moving Average). El cual deriva de sus tres componentes: Autorregresivo (AR), Integrado (I) de medias

móviles (MA); cuyo objetivo es identificar y estimar un modelo estadístico que pueda ser interpretado como generador de información muestra.

### 2.1.3.1. Autorregresivo AR(p)

Describe una clase particular de proceso en que las observaciones en un momento dado son predecibles a partir de las observaciones previas del proceso más un término de error. El caso mas simple es de ARIMA (1,0,0) o AR(1) o del primer orden cuya expresión matemática es:

$$Y_t = \phi_1 Y_{t-1} + a_t \quad (2.1)$$

Un modelo Autorregresivo de orden p, toma la forma

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (2.2)$$

que mediante el operador de cambio retroactivo B, en la forma:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = a_t \quad B^k (Y_t) = Y_{t-k} \quad (2.3)$$

que es estacionario si las raíces del polinomio en B dado por  $1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  caen fuera del círculo unidad. Esa condición es equivalente a que las raíces de la ecuación:

$$y^p - \phi_1 y^{p-1} - \phi_2 y^{p-2} - \dots - \phi_{p-1} y - \phi = 0 \text{ Sean todas inferiores a uno en módulo.}$$

Un proceso autorregresivo siempre es irreversible.

### 2.1.3.2. Modelo de media móvil MA(q)

Describe una serie temporal estacionaria. El valor actual puede predecirse a partir de la componente aleatoria de este momento y, en menor medida, de los impulsos aleatorios anteriores.

El modelo ARIMA (0,0,1) también denotado por MA(1), viene dado por la expresión:

$$Y_t = a_t - \theta_1 a_{t-1} \quad (2.4)$$

El proceso de medias móviles de orden q viene dado por la expresión:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.5)$$

Que de igual forma que el modelo anterior puede ponerse en la forma:

$$Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.6)$$

Un proceso de medias móviles es siempre estacionario, además es invertible si las raíces del polinomio en B definido por  $1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$  caen fuera del círculo unidad. Esta condición es equivalente a que las raíces de la ecuación  $y^q - \phi_1 y^{q-1} - \phi_2 y^{q-2} - \dots - \phi_{q-1} y - \phi_q = 0$  sean todas inferiores a uno en módulo.

### 2.1.3.3. Modelo Autoregresivo de media móvil ARMA(p,q)

Una extensión natural de los modelos AR (p) y MA (q) es un tipo de modelos que incluyen tanto términos autorregresivos como de medias móviles y se

definen como ARMA (p,q) o también como ARIMA(p,0,q), y está representado por:

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.7)$$

puede ponerse de la forma:

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.8)$$

o sea:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.9)$$

El proceso ARMA (p,q) es *estacionario* si lo es su componente autorregresiva, y es *invertible* si lo es su componente de media móviles.

#### **2.1.3.4. Modelo Autoregresivo integrado de media móvil ARIMA(p,d,q)**

Éste modelo es una serie temporal que se convierte en un ruido blanco (proceso puramente aleatorio) después de ser diferenciada  $d$  veces. El modelo ARIMA (0,d,0) se expresa mediante  $(1-B)^d Y_t = a_t$  y toma la siguiente expresión:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) (1 - B)^d Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.10)$$

Permite describir una serie de observaciones después de que hayan sido diferenciadas  $d$  veces a fin de extraer las posibles fuentes de no estacionariedad. Ésta fórmula general se puede aplicar a cualquier modelo. Si hay alguna componente  $p, d, q$  igual a cero, se elimina el término correspondiente de la fórmula general. Los modelos cíclicos o estacionales son

aquellos que se caracterizan por oscilaciones cíclicas a veces se superponen a una tendencia secular. La serie con tendencia secular y variaciones cíclicas pueden representarse mediante los modelos ARIMA (p,d,q) (P,D,Q). El primer paréntesis se refiere a la tendencia secular o ate regular de la serie y el segundo paréntesis se refiere a las variaciones estacionales, o parte cíclica de la serie temporal.

La expresión de un modelo ARIMA (0,1,1)(0,0,1)<sub>12</sub> será:

$$(1 - B)Y_t = (1 - \theta_1 B^{12})(1 - \delta_{12} B^{12}) \quad (2.11)$$

La expresión de un modelo ARIMA (0,1,1)(0,1,1)<sub>12</sub> será:

$$(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)(1 - \delta_{12} B^{12})a_t \quad (2.12)$$

La expresión de un modelo ARIMA (2,1,0)(1,0,0)<sub>12</sub> será:

$$(1 - \phi_1 B^2 - \phi_2 B^{12})(1 - \Omega_1 B^{12})(1 - B)Y_t = a_t \quad (2.13)$$

La expresión de un modelo ARIMA (1,1,1)(2,1,1)<sub>12</sub> será:

$$(1 - \phi_1 B)(1 - \Omega_1 B^{12} - \Omega_2 B^{24})(1 - B^{12})(1 - B)Y_t = (1 - \theta_1 B)(1 - \delta_{12} B^{12})a_t \quad (2.14)$$

Si este modelo estimado se utilizará para predicción lo primero que surge es ¿Cómo sabe si surge un proceso AR puro, (de ser así cuál es el valor de p) o un proceso MA puro, (de ser así cuál es el valor de q) o un proceso ARMA (de ser así, cuales son los valores de p y q) o un proceso ARIMA, en cuyo caso se deben conocer los valores p, d y q?

Determinar el tipo de modelo más adecuado para la serie. La decisión se tomará en base de los correlogramas de las funciones de autocorrelación (ACF) y de autocorrelación parcial (PACF) que corresponden a los procesos autoregresivos y de medias móviles de las componentes regular y estacional. En caso de duda puede considerarse varias alternativas que, luego de contrastarse su validez se define el modelo adecuado. Para ello podemos hacer uso de la siguiente tabla:

**TABLA 1.- Patrones Teóricos de ACF Y PACF**

	<b>ACF</b>	<b>PACF</b>
<b>AP(p)</b>	Decrecimiento rápido de tipo geométrico puro, y geométrico con alternancia de signos, sinusoidal o mezcla de varios tipos.	Se anula para retardos superiores a p
<b>MA(q)</b>	Se anula para retardos superiores a q.	Decrecimiento rápido de tipo exponencial y/o sinusoidal
<b>ARMA(p,q)</b>	Los primeros valores iniciales no tienen patrón fijo y van seguidos de una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas	Los primeros valores iniciales no tienen patrón fijo y van seguidos de una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas.

#### **2.1.4. SOBREAJUSTE Y SELECCIÓN DEL MODELO**

El sobre ajuste consiste en estimar cada uno de los modelos obtenidos al aumentar en una unidad cada uno de los órdenes p, q, P y Q. Una vez sobre ajustado los parámetros p, q, P y Q se decide sobre el modelo definitivo.

### **2.1.5. VERIFICACIÓN DEL DIAGNÓSTICO**

Se usan procedimientos para evaluar el o los modelos preseleccionados, contraste de significación de parámetros, covarianza entre estimadores, coeficiente de correlación, suma de cuadrado de errores, etc.

Para ajustar el modelo se analizan los residuos que se obtienen del ruido blanco ( $a_t$ ). Un proceso estocástico es ruido blanco si:

- La esperanza de  $a_t$  es igual a cero
- La varianza de  $a_t$  es constante
- La correlación de los  $a_t$  es igual a cero
- $a_t$  sigue una distribución normal

### **2.1.6. PREDICCIÓN**

Una vez seleccionado el modelo, se hace la predicción.

## **2.2. TEORÍA DEL CAOS**

### **2.2.1. CARACTERIZACIÓN DEL CAOS**

La Teoría del Caos permite deducir el orden subyacente que ocultan fenómenos aparentemente aleatorios. Se sabe que ecuaciones totalmente

deterministas (como el set de Lorenz) presentan las siguientes características que definen el Caos:

**a. Son deterministas**, es decir: Existe una "ley" que gobierna la conducta del sistema (¿Qué es lo contrario de "determinista"? ¿"Aleatorio" o "con libre albedrío"? ¿Existe el Libre Albedrío para las Ciencias Duras o es sólo una ilusión?)

- El fenómeno se puede expresar por "comprensión" en lugar de hacerlo por "extensión"
- Existe una simulación de menor tamaño (Kb) que el sistema original que permite generar los mismos datos observados. Cabe señalar que según Chaitin (1994) un sistema es aleatorio cuando el algoritmo que genera su propia serie ocupa más Kb que el sistema original (por lo tanto, lo más eficiente es expresar el sistema por "extensión" y no por medio de un algoritmo)

**b. Son muy sensibles a las condiciones iniciales:**

- Una desviación infinitesimal en el punto de inicio provoca una divergencia exponencial en la trayectoria del Espacio de Fase, lo que se puede cuantificar con el "Exponente de Lyapunov"
- La extrema sensibilidad a las condiciones iniciales implica que el comportamiento del sistema se indetermina a partir de cierto "Horizonte de Predicitibilidad", dado que la incerteza tecnológica

asociada a los datos de entrada siempre va a ser mayor que el concepto de "infinitesimal matemático"

- A pesar de la impredecibilidad de una trayectoria particular del Espacio de Fase, se pueden encontrar "Atractores" o zonas del Espacio de Fase que tienden a ser "visitadas" con mayor frecuencia que otras.

NOTA: Normalmente la trayectoria en el Espacio de Fase de un sistema caótico genera una curva fractal (de dimensión fraccionaria)

**c. Parecen desordenados o aleatorios, pero en el fondo no lo son:**

- Siguen ecuaciones deterministas
- Presentan Atractores

Un ejemplo de ecuación determinista pero caótica es:

$$Y_t = 4Y_{t-1}(1 - Y_{t-1}) \quad (2.15)$$

Se puede ilustrar el Efecto Mariposa comparando los gráficos que resultan cuando se utilizan las siguientes condiciones iniciales:

- Sistema A:  $X_0 = 0.399999$
- Sistema A + una mariposa:  $X_0 = 0.400000$  (apenas una millonésima de diferencia)

### **2.2.2. SISTEMAS CAÓTICOS**

En los sistemas reales es frecuente encontrar señales que aparentemente tienen un comportamiento casual, caracterizado por una elevada sensibilidad a las condiciones iniciales e imprevisibilidad a través del tiempo.

Este tipo de sistemas se los define normalmente como caóticos, los cuales pueden ser caracterizados a través de ciertas variables en el espacio de las fases donde las dimensiones representan variables dinámicas.

### **2.2.3. CARACTERÍSTICAS GENERALES DE LOS SISTEMAS CAÓTICOS**

Generalmente un sistema dinámico está descrito por una ecuación diferencial, pero es conocido que este tipo de ecuaciones se puede presentar en términos de un mapa (Poincaré).

Lo cual es muy útil para visualizar las características cualitativas de un comportamiento caótico o no caótico, así podemos reducir una ecuación a un mapa de Poincaré. De esta manera, es fácil comprender los varios comportamientos de un sistema:

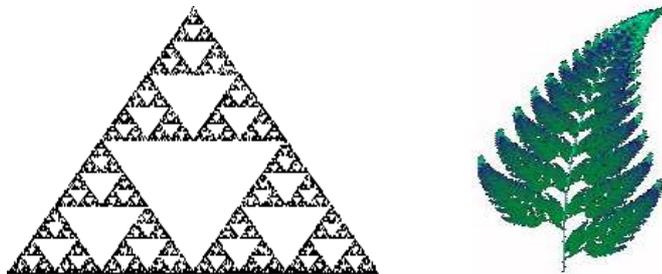
- a) Si la trayectoria del sistema es periódica, sobre la sección o mapa de Poincaré se tiene un conjunto finito de puntos aislados.

- b) Si el movimiento es casi periódico, se tendrá una figura regular cerrada.
- c) Si es un movimiento caótico, formara una mancha no estructurada.

El paso de la ecuación diferencial al mapa no cambia la naturaleza regular o caótica del sistema estudiado.

#### **2.2.4. GEOMETRÍA FRACTAL**

La geometría fractal es sin duda la geometría de la naturaleza: las nubes, nuestro sistema circulatorio, los cauces de grandes ríos, las cadenas montañosas, etc.



**GRAFICO 6.-** Triángulo de Sierpinski. Hoja de Helecho.

La importancia de la geometría fractal como apoyo al estudio de la complejidad radica según Cambell (1984) en cuatro puntos principales:

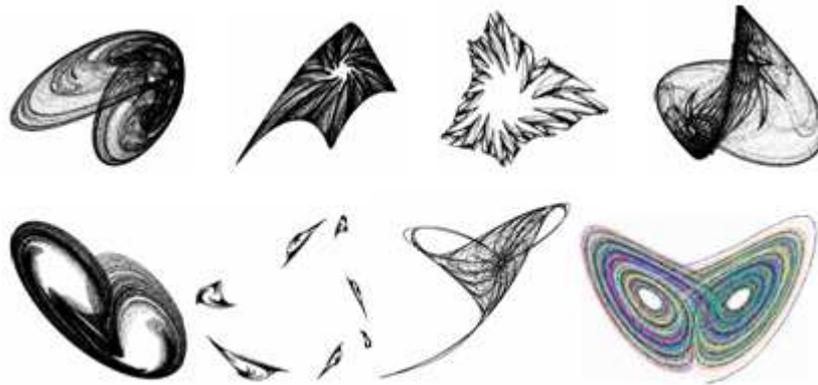
- a) Provee dimensiones adicionales y más cercanas a la realidad en comparación con la geometría Euclidiana

- b) La mayoría de los sistemas complejos son caóticos, y estos exhiben conductas extrañas asociadas con límites o campos que no pueden ser representados en dimensiones enteras.
  
- c) Los sistemas dinámicos pueden ser representados en series de tiempo y sus dimensiones son importantes si se busca estudiarlos.
  
- d) Los fractales son escalables, esto es, se puede reducir o ampliar su análisis para observar detalles, mientras que las formas básicas se conservan.

El paso de la ecuación diferencial al mapa no cambia la naturaleza regular o caótica del sistema estudiado.

### **2.2.5. ATRACTOR**

El atractor de un sistema dinámico es la región del espacio de las fases, en la cual la trayectoria es atraída, tendiendo el movimiento luego de un tiempo de ajuste al atractor.



**GRAFICO 7.-** Representación visual de atractor

Una vasta clase de sistemas dinámicos disipativos, tienen la notable propiedad de poseer un atractor, el cual es el lugar de los puntos recorridos luego de una trayectoria suficientemente larga. Un ejemplo simple es el del péndulo oscilante en el aire, la amplitud de oscilación se atenúa progresivamente y todas las trayectorias independientemente de las condiciones iniciales se acercan al punto fijo estable que constituye el atractor.

En el caso de los sistemas caóticos, en las cuales se tiene una fuerte dependencia de las coordenadas iniciales, el atractor no es en efecto un objeto regular, sino mas bien tiene una estructura muy complicada, entonces se lo llama atractor extraño (Ruelle), el cual puede tener una estructura fractal.

Para tener un atractor es necesario tener una contracción del volumen del espacio de las fases, excepto para ciertos casos (movimientos sin rozamiento, astronómicos, etc.). Para un sistema caótico se tiene un comportamiento diverso, por ejemplo, para el atractor de Lorenz, el volumen tiende a transformarse en una cinta, de espesor siempre más pequeña, con ancho al

iniciar constante y largo creciente exponencialmente. Al pasar del tiempo la cinta se repliega sobre si misma y se envuelve sobre el atractor. A este tipo de atractores se los llama extraños.

### 2.2.6. DIMENSION FRACTAL

En general se dice que un objeto de dimensión  $d$  puede ser puesto en correspondencia biunívoca con  $d$  números reales de modo que cambiando poco la posición del punto este número cambia poco. La dimensión así definida es el número de direcciones independientes accesibles a un punto que se mueve sobre el objeto, a la cual se le llama dimensión topológica y es evidentemente un número entero no negativo inferior o igual a la dimensión del espacio en el cual el objeto es puesto.

Pero esta dimensión no esta en grado de describir completamente las características geométricas de un objeto, por lo que es necesario introducir un nuevo índice que logre distinguir el objeto con la misma dimensión topológica, llamada dimensión fractal (Mandelbrot, 1987): un objeto tiene dimensión fractal  $D_f$  si el número  $N(l)$  de cubitos de lado  $l$  se comporta en el límite pequeño  $l$  del siguiente modo.

$$N(l) \approx \frac{1}{l^{D_f}} \quad (2.16)$$

Para calcular la dimensión fractal, se puede considerar un mapa cuya ley de evolución es dada por el mapa  $x(1), x(2), \dots, x(M)$   $x(l)$ , con  $M \gg 1$ , se divide luego en el espacio de las fases en elementos de lado 1, de modo que al menos contengan uno de los puntos de la serie, obteniendo así,  $N(l)$ , se repiten para diversos valores de 1, obteniendo un gráfico de  $\ln(N(l))$  en función de  $\ln(l)$ , obteniendo aproximadamente una recta cuya pendiente es  $D_f$  aunque este método es conceptualmente simple en la práctica calcularlo resulta complejo, especialmente por la excesiva memoria que ocupan los programas que lo calculan.

Afortunadamente existe un método debido a Grassberger y Procaccia (1983), que permite calcular con buena aproximación a través de una función de correlación la dimensión fractal  $D_f$ .

$$C(l) = \frac{2}{M(M-1)} \sum_{i,j} \theta(l - |x(i) - x(j)|) \approx l^{\nu} \quad (2.17)$$

Con  $X(i)$  y  $X(j)$  número de puntos pares con distancia menor que  $l$  de una sucesión  $x(1), x(2), \dots, x(M)$ , que en el caso de ecuaciones diferenciales se sustituye por  $x(\tau), x(2\tau), \dots, x(m\tau)$ , donde  $t$  es el intervalo de correlación (pariamiento), que conviene no tomarlo muy grande, donde  $\tau$  se llama tiempo de retardo y  $m$  dimensión de encaje.

Si el objeto es regular el  $v$  es entero y coincide con  $D_f$ , mientras si la distribución no es uniforme el  $v < D_f$ . Pero se puede demostrar que, estos coinciden cuando, se cumple

$$D_F \approx v = \lim_{l \rightarrow 0} \frac{C(l)}{\log l} \quad (2.18)$$

### **2.2.7. TIEMPO DE RETARDO**

Es el paso más importante en el análisis del espacio de fases, es la técnica de reconstrucción del sistema, formando vectores y ejecutando una proyección en el espacio de fases, que es un espacio físico que permite detectar variables dinámicas.

### **2.2.8. DIMENSIÓN DE ENCAJE**

Aquí se determinan las dimensiones apropiadas para las matrices creadas, buscando vecinos a un punto fijo y que estén contenidos dentro de una bola de radio centrada en ese punto, de esta forma se calcula la distancia interactiva y se computa. Si la relación de la distancia de la interacción y del vecino mas cercano excede a un valor determinado, el punto se marca como un vecino equivocado.

Determinar la dimensión del espacio de fases es decisivo en el objetivo para desdoblar las órbitas del atractor y evitar las auto intersecciones que podríamos encontrar en espacios de menores dimensiones, estudiar la relación de vecindad entre puntos en el espacio de fases para distintas dimensiones sirve a dicho propósito.

El **teorema de Takens**, dice que si la dimensión del atractor es  $x$  número entero, entonces para gráficos genéricos retrasados, la dimensión del embedding es al menos dos veces el número entero más uno. Esta técnica es una de las pocas además, que se han mostrado útiles para el estudio de escenarios dinámicos caóticos y poder hacer predicciones a corto plazo de su evolución.

### 2.2.9. EXPONENTES DE LYAPUNOV

El caos surge del crecimiento exponencial de las perturbaciones infinitesimales, para garantizar el análisis de esta inestabilidad están los exponentes de Lyapunov, que cuantifican cuando dos trayectorias se alejan.

$$\lambda = \frac{1}{t} \ln \left( \frac{l(t)}{r} \right) \quad (2.19)$$

Con  $l(t)$  semieje mayor de una elipse,  $r$  radio pequeño y  $t$  tiempo grande, en general se establece:

$$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_d$$

Variables que permiten caracterizar un atractor, así:

- a) Para un punto fijo todos los  $\lambda_i$  son negativos.
- b) En un ciclo límite  $\lambda_1 = 0$  y  $\lambda_i < 0$  por  $i > 1$ .
- c) En el movimiento de un toro n dimensional,  $\lambda_1 = \lambda_2 = \dots = \lambda_n$  y  $\lambda_i < 0$  por  $i > n$ .
- d) En un sistema caótico al menos un exponente de Lyapunov es negativo.

A partir de este resultado se puede definir, la entropía de Kolmogorov – Sinai.

$$h = \sum_{i=\#\lambda>0} \lambda_i \quad (2.20)$$

O la dimensión de Kaplan – Yorke (1979)

$$D_F = j + \frac{\sum_{i=1}^j \lambda_i}{|\lambda_{j+1}|}, \text{ con } \sum_{i=1}^j \lambda_i < 0 \quad (2.21)$$

Con j coeficientes positivos de Lyapunov.

# **CAPÍTULO III**

## **RESULTADOS**

### **3. RESULTADOS**

#### **3.1. MODELO ARIMA DE BOX-JENKINS**

Se incorporan los datos disponibles (promedios diarios de 7 años a partir del 2000 asta el 2006) como un vector, es decir en una única columna con el nombre  $Y = \text{"caudal"}$  (L/s.)

Los datos que analizaremos son los de caudales medios diarios de río Blanco medidos por en el Simatic S5 en la central hidroeléctrica de Quimiag; mismo que son copiados manualmente en un formato estándar de control con varios datos más; luego la documentación es trasladada a las oficinas para digitar sus resúmenes en la estación uno de la Empresa Eléctrica Riobamba S.A. (EERSA) y son almacenados en el departamento de despacho de la estación dos de ésta Institución.

Para lo anterior contamos con una serie de tiempo de los caudales medios diarios que parte en el mes de enero del 2000 y termina en diciembre de 2006, es decir 2 558 observaciones.

Dejaremos los últimos datos de la serie como periodo de validación, es decir el último mes quedará reservado para efectos de comprobación de la capacidad predictiva del modelo.

A continuación se entrega un cuadro con las estadísticas básicas correspondientes a 2 558 observaciones.

**Tabla 2.-** Estadísticas descriptivas de los caudales medios diarios de río Blanco (2000-2006)

**Estadísticos**

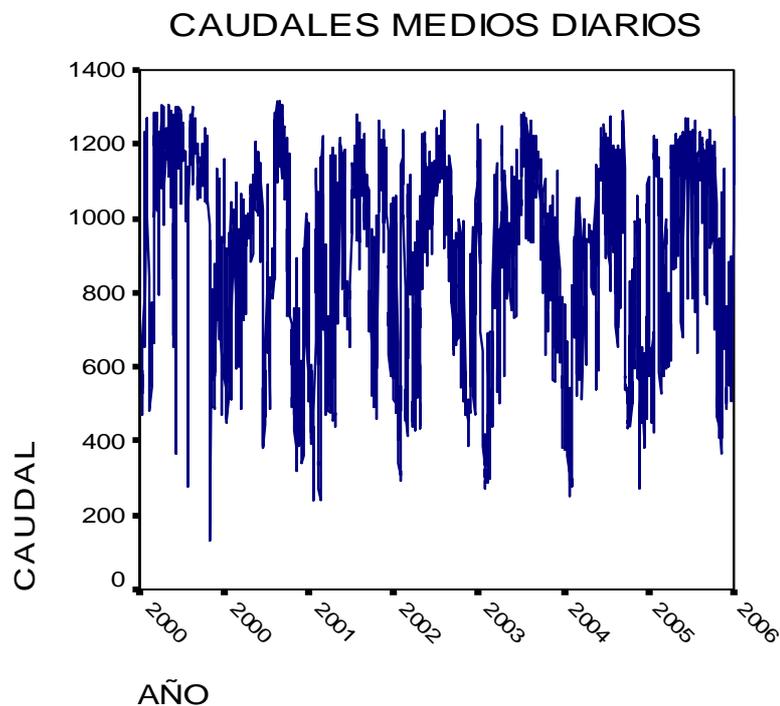
CAUDAL

N	Válidos	2527
	Perdidos	0
Media		887,77
Mediana		939,00
Moda		972,00 <sup>a</sup>
Desv. típ.		265,56
Varianza		70522,61
Rango		1188,00
Mínimo		129,00
Máximo		1317,00

a. Existen varias modas. Se mostrará el menor de los valores.

Podemos ver que diariamente, en promedio se tiene un caudal de 887,77 L/s el caudal más bajo corresponde a 129,00 L/s enero 19 del 2002 y la más alta de 1317 L/s en agosto 11 del 2001.

### 3.1.1. GRÁFICA DE LA SERIE



*Fuente:* Estación N° 2 EERSA.

**GRAFICO 8.-** Serie temporal de caudales de río Blanco. Período 2000 - 2006

### 3.1.2. TRANSFORMACIONES

La primera etapa en el proceso de modelamiento o identificación del modelo univariante de una serie temporal, consiste en comprobar si esta es estacionaria en media y en varianza. Si además los datos de la serie están recogidos con una periodicidad inferior a un año (datos de los promedios diarios) tendremos que comprobar también si es estacionaria estacionalmente. En el caso que la serie no sea estacionaria, ya sea en varianza o en media, habrá que aplicarle las transformaciones necesarias para que lo sea.

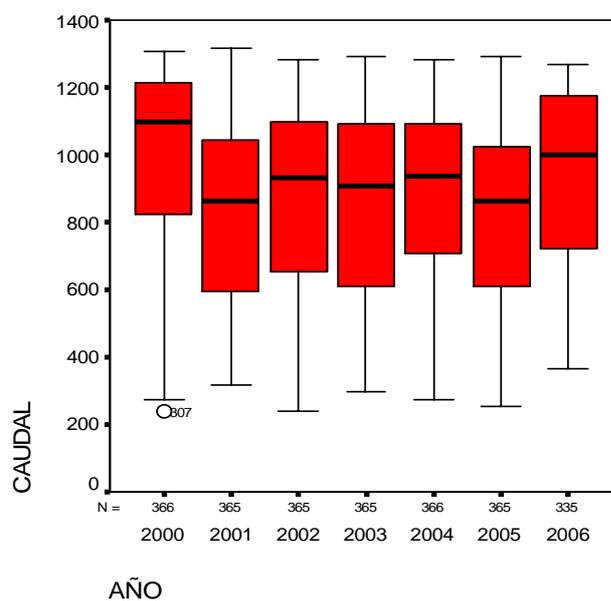
Para detectar si la varianza de la serie es constante, es decir si no cambia con el tiempo, aplicaremos la prueba de Levene para la homogeneidad de varianzas. Dado que el periodo de ajuste comprende desde 2 000 hasta 2 006. La prueba de Levene permitirá contrastar la hipótesis nula de que no existen diferencias significativas entre las varianzas de la serie. A continuación se muestra una tabla con los resultados y después un Box Plot .

**Tabla 3.-** Prueba de Levane para la homogeneidad de varianza

Prueba de homogeneidad de varianzas

CAUDAL

Estadístico de Levene	gl1	gl2	Sig.
,854	6	2520	,528

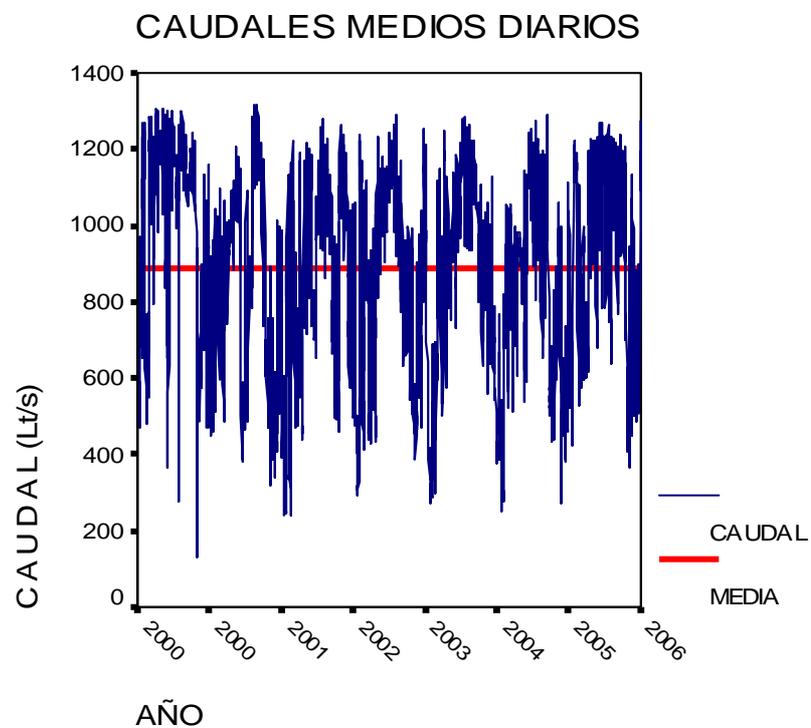


**GRAFICO 9.-** Diagrama de Box Plot

Utilizando la Prueba de Homogeneidad de Varianza de Levene, llegamos a la conclusión que no hay suficiente evidencia para rechazar la hipótesis nula de

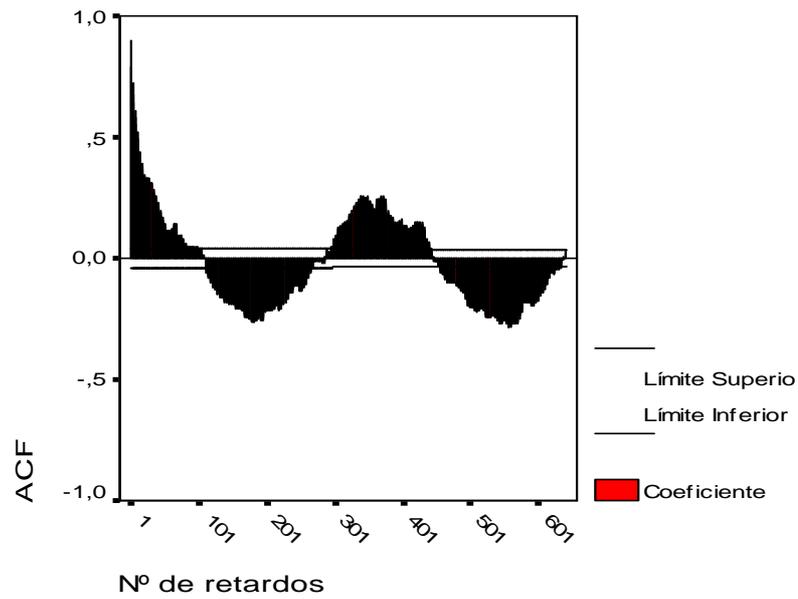
que las varianzas son las mismas a lo largo de los años al nivel de significancia 0,05. A la misma conclusión se llega observando el gráfico 9, lo que quiere decir que la serie es estacionaria en varianza.

La serie de datos también tiene que ser estacionaria en media, para comprobar lo anterior veremos algunos gráficos.



**GRAFICO 10.-** Visualización de los caudales medios diarios y su media

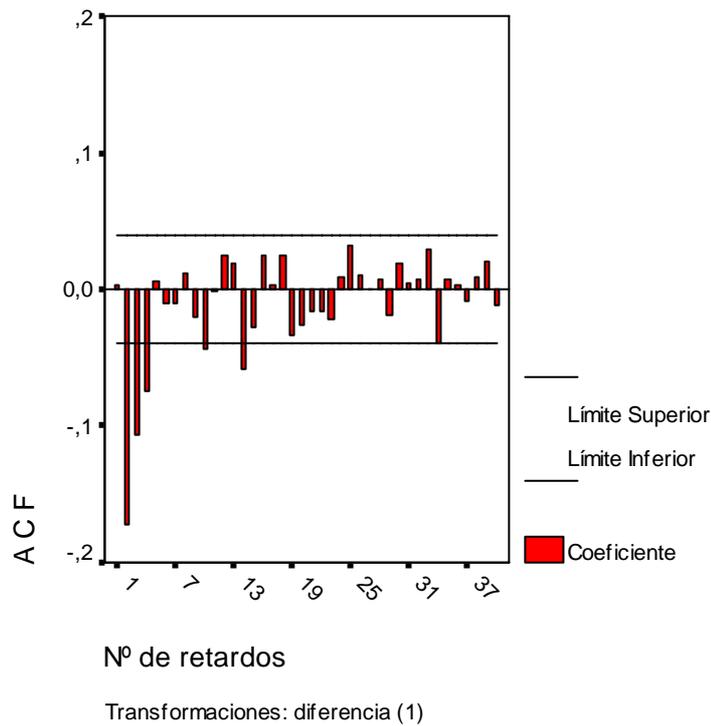
Observe que hay algunos valores por debajo de la media y otros por sobre la media, lo que indica no estacionariedad en media, la media es la línea roja.



**GRAFICO 11.-** Correlograma simple de la serie con los 640 primeros retardos.

El gráfico anterior nos muestra el correlograma de la serie, específicamente la función de autocorrelación. Nótese como los picos van descendiendo paulatinamente a cero, los primeros rezagos de la serie están fuera de los límites de confianza, otra evidencia de que la serie no es estacionaria en media.

Para lograr la estacionariedad en media se procederá a diferenciar una vez la serie y para poder visualizar de mejor manera analizaremos el correlograma con los primeros 40 retardos.



**GRAFICO 12.-** Autocorrelación simple de los primeros 40 rezagos o retardos

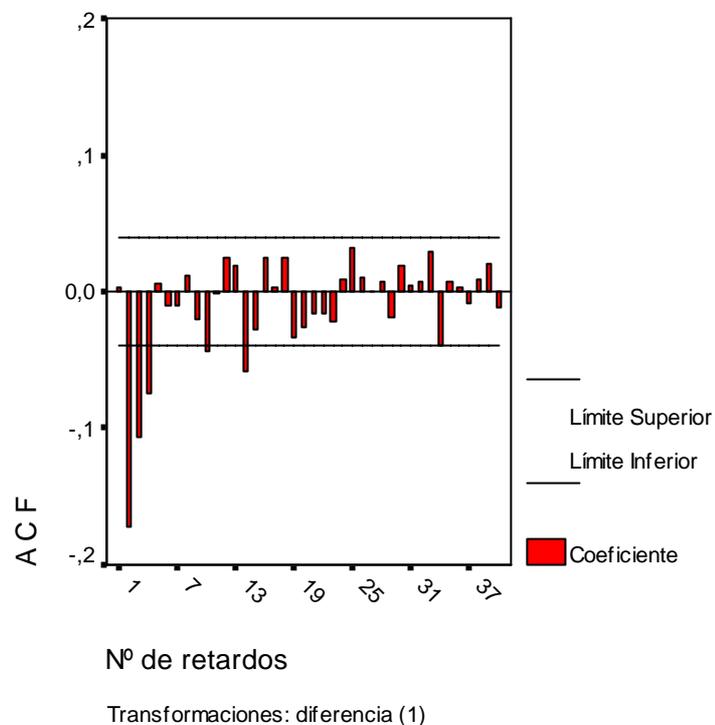
Ahora se observa en el gráfico 12 de autocorrelación simple que a partir de los tres primeros retardos la serie no se va reduciendo paulatinamente a cero. Además se aprecia claramente que la serie es estacionaria estacionalmente, por lo tanto no será necesaria hacer una diferenciación estacional.

Esperamos que después de todas estas transformaciones, llámense aplicación de logaritmos naturales y diferenciaciones, se muestre la estructura o la caja negra que genera la serie.

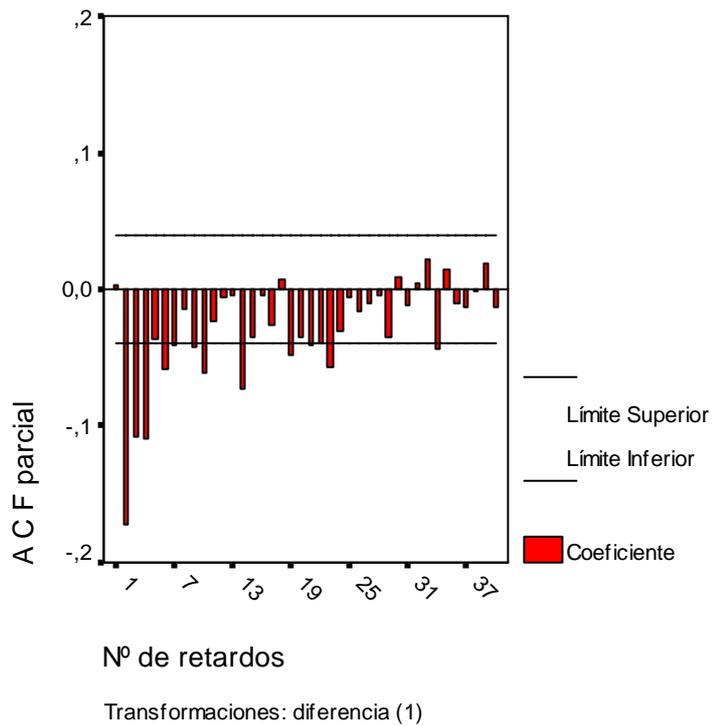
Se recuerda que la serie para hacerla estacionaria en media, fue necesario hacerle una diferenciación a la parte no estacional. Esa es la serie que se seguirá aplicando desde ahora y para la cual buscaremos su identificación.

### 3.1.3. IDENTIFICACIÓN DEL MODELO

Necesitamos 2 tipos de gráficos, la función de autocorrelación (ACF ) de nos muestra la presencia o no de las medias móviles del modelo MA(p) y la función de autocorrelación parcial (PACF), que nos muestra la posible presencia de autocorrelaciones AR(q)



**GRAFICO 13.-** Función de Autocorrelación simple sobre los primeros 40 retardos de río Blanco 2000-2006



**GRAFICO 14.-** Función de autocorrelación parcial sobre los primeros 40 retardos de río Blanco 2000-2006

Observando los gráficos anteriores se aprecia que tanto la función de autocorrelación simple y la función de autocorrelación parcial descienden paulatinamente a cero, lo que podría estar señalando un proceso de generación AR (1), MA (1).

Por otra parte, observando la parte estacional de la serie, en el gráfico de autocorrelación parcial, se aprecia que los retardos caen paulatinamente a cero, al menos se observa eso en el retardo 7, en cambio en la función de autocorrelación el único retardo que cae fuera de los límites de confianza es el número 12 y después siempre sus múltiplos se encuentran dentro de los límites

de confianza. Lo anterior nos dice que el proceso más apropiado es un MA (1) para la parte estacional de la serie.

#### **3.1.4. SOBREAJUSTE Y SELECCIÓN DEL MODELO**

El modelo ARIMA a utilizar para la estimación es el modelo ARIMA (1,1,1) x (1,0,1)<sub>12</sub>

La nomenclatura anterior es una de las maneras de identificación de modelos ARIMA. El primer paréntesis señala la estructura no estacional de la serie, en particular señala que a la serie se le ha efectuado una diferenciación o se ha integrado, para hacerla estacionaria (de ahí viene la I de ARIMA). En el segundo paréntesis se denota la parte estacional de la serie. Se debe señalar que no todos los modelos ARIMA tienen un segundo paréntesis, este lo tiene por que utiliza datos menores al año y por que éstos, al parecer, son parte de la caja negra que genera la serie y que estamos interesados en identificar. Volviendo a nuestra descripción, la parte estacional, como es estacionaria, por eso no se tuvo que diferenciar y también el retardo del modelo es MA (1), específicamente SMA (1) (la S es por estacionalidad).

**Tabla 4.- Estadísticos del modelo ARIMA(1,1,1)(1,0,1)<sub>12</sub>**

	<b>B</b>	<b>SEB</b>	<b>T-RATIO</b>	<b>APPROX. PROB.</b>
<b>AR1</b>	-0,74452026	0,1095111	-6,798582	0,00000000
<b>MA1</b>	-0,79570777	0,0997193	-7,9794746	0,00000000
<b>SAR1</b>	0,5803895	0,2889178	2,0088393	0,04466061
<b>SMA1</b>	0,61738824	0,2787105	2,2151598	0,02683817
<b>CONSTANT</b>	-0,11739356	2,2544336	-0,0520723	0,9584752

La tabla 4 muestra los estadísticos más importantes del modelo, lo que nos interesa en estos modelos es que sus regresores sean estadísticamente significativos, y eso nos lo señala la última columna. Los p-valores asociados a cada uno de los regresores son significativos al 0,05.

El hecho de que el modelo identificado sea adecuado no debe descartar la posibilidad de que otro modelo algo más complejo pueda ajustarse mejor a la serie observada. Para comprobar si algún modelo se ajusta mejor conviene utilizar la técnica del sobreajuste.

Por lo tanto, si el modelo candidato recién calculado es:

ARIMA (1,1,1) x (1,0,1)<sub>12</sub>

Los sobreajustes serán:

1. ARIMA (1,1,2) x (1,0,1)<sub>12</sub>
2. ARIMA (1,1,2) x (2,0,1)<sub>12</sub>
3. ARIMA (2,1,1) x (2,0,1)<sub>12</sub>

#### 4. ARIMA (2,1,1) x (1,0,1)<sub>12</sub>

Después de hacer los cálculos, se concluye que: el regresor AR(1) del modelo 1 no era significativo, el regresor del modelo 2 no era significativo, los regresores SMA(1) y SMA(2) del modelo 3 no eran significativos; por lo tanto los modelo 1,2 y 3 fueron rechazados. El modelo 4 sí tuvo un buen resultado.

**Tabla 5.-** Datos Modelo 4 ARIMA (2,1,1)(1,0,1)<sub>12</sub>

	<b>B</b>	<b>SEB</b>	<b>T-RATIO</b>	<b>APPROX. PROB.</b>
<b>AR1</b>	0,83506463	0,02656387	31,436104	0,00000000
<b>AR2</b>	-0,12944728	0,02120568	-6,104368	0,00000000
<b>MA1</b>	0,89369544	0,01902997	46,962533	0,00000000
<b>SAR1</b>	0,79263354	0,20965675	3,780625	0,00016007
<b>SMA1</b>	0,81439344	0,1988164	4,096208	0,00004332
<b>CONSTANT</b>	-0,02223317	0,75542553	-0,029431	0,97652292

Como podemos observar los p-valores asociados a cada uno de los regresores es cero, en otras palabras a un 0,05 nivel de significancia no se puede rechazar la hipótesis nula de que son cero.

Ya tenemos 2 modelos candidatos, ahora veremos cual de los 2 es mejor. Para saberlo nos fijaremos en cómo los modelos se ajustan a los datos al periodo de ajuste (enero de 2000 a noviembre de 2006) y al periodo de validación (diciembre del 2006)

Si uno sólo se fija en el periodo de ajuste el segundo modelo, presenta un mejor ajuste, sin embargo para compararlos se hará la prueba de Levene de la homogeneidad de varianza para los residuos de cada uno de los modelos.

**Tabla 6.-** Prueba de Levene de homogeneidad de varianza de los residuos o errores ARIMA (1,1,1)(1,0,1)<sub>12</sub> y ARIMA (2,1,1)(1,0,1)<sub>12</sub>

**Prueba de homogeneidad de varianzas**

	Estadístico de Levene	gl1	gl2	Sig.
Error for CAUDAL from ARIMA, (1,1,1)(1,0,1)	2,649	6	2519	,015
Error for CAUDAL from ARIMA, (2,1,1)(1,0,1)	3,217	6	2519	,004

La prueba de Levene muestra que no hay motivos suficiente para rechazar la hipótesis nula de homogeneidad de varianza al 0.01 de nivel de significancia el modelo ARIMA(1,1,1)(1,0,1)<sub>12</sub>, mientras que para el modelo ARIMA (2,1,1)(1,0,1)<sub>12</sub> se rechaza la hipótesis nula, quiere decir, no hay homogeneidad de varianza, por lo tanto este modelo no es el mas apropiado debido que no cumple con las condiciones de ruido blanco.

Corresponde ahora calcular los parámetros que respondan a las ecuaciones del modelo. En la tabla 7 se observa la estimación de los parámetros del modelo seleccionado, indicándose en la columna B los valores para cada uno de los coeficientes del modelo y en la última columna la significación de las estimaciones.

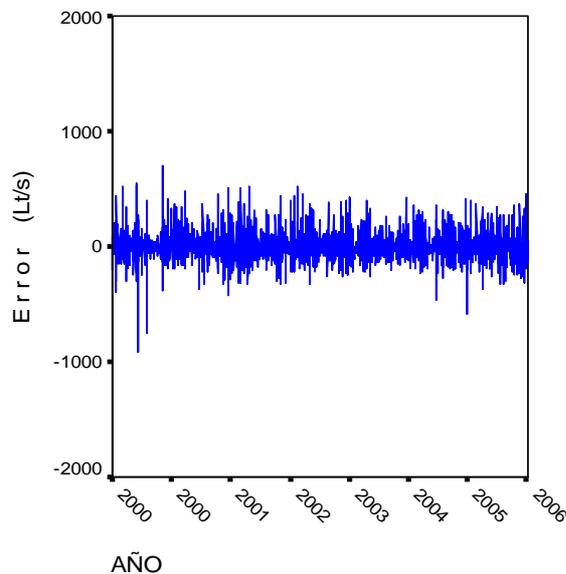
La constante calculada en el modelo tiene una significación cercana a uno, por lo que puede eliminarse y realizar el ajuste del modelo sin considerarla, obteniéndose los valores que se presentan en la Tabla 7, en el cual pueden observarse leves modificaciones en los valores de la variable.

**Tabla 7.-** Ajuste del modelo ARIMA (1,1,1)(1,0,1)<sub>12</sub>

	<b>B</b>	<b>SEB</b>	<b>T-RATIO</b>	<b>APPROX. PROB.</b>
<b>AR1</b>	-0,72058981	0,11885176	-6,0629292	0,00000000
<b>MA1</b>	-0,77364556	0,10900269	-7,0974904	0,00000000
<b>SAR1</b>	0,65292744	0,25284446	2,5823284	0,00986951
<b>SMA1</b>	0,68668207	0,24228533	2,8341876	0,00463079

### 3.1.5. VERIFICACIÓN DEL DIAGNÓSTICO

Se efectúa luego la validación del modelo utilizando los errores que se obtienen como diferencia entre los valores de la serie y la resultante del modelo.



**GRAFICO 15.-** Errores que se obtienen como diferencia entre los valores de la serie de caudales originales y la resultante del modelo.

En este paso se trata de ver si el modelo seleccionado se ajusta a los datos en forma razonable. Una simple prueba del modelo seleccionado es ver si los residuales estimados a partir de este modelo son de ruido blanco; si lo son puede aceptarse el ajuste particular, caso contrario se debe empezar nuevamente por lo tanto la metodología Box-Jenkins es un proceso iterativo.

**Tabla 8.- Estadísticas Descriptivas de los Residuos**

**Estadísticos**

Error for CAUDAL from ARIMA, MOD\_1 NOCON

N	Válidos	2526
	Perdidos	1
Media		-,1311438
Error típ. de la media		2,398560
Mediana		-9,19353
Moda		-909,381 <sup>a</sup>
Desv. típ.		120,5500
Varianza		14532,31
Rango		1612,436
Mínimo		-909,381
Máximo		703,05496
Suma		-331,269

a. Existen varias modas. Se mostrará el menor de los valores.

Se puede ver que la media de los residuos es de -0,1311, además se muestran algunas estadísticas de la misma variable, lo más rescatable que si se analiza si el supuesto de que el residuo del modelo presenta media igual a cero se cumple, para ello hacemos una prueba de hipótesis suponiendo media de los residuos igual a cero.

**Tabla 9.-** Prueba cero para la comparación de la media de los residuos

**Prueba para una muestra**

	Valor de prueba = 0					
	t	gl	Sig. (bilateral)	Diferencia de medias	95% Intervalo de confianza para la diferencia	
					Inferior	Superior
Error for CAUDAL from ARIMA, MOD_9 NOCON	-,055	2525	,956	-,1311438	-4,83449	4,5722027

Como se observa en la tabla 9, no hay suficiente evidencia para rechazar la hipótesis nula, es decir al 95% de confianza la media de los residuos es cero.

**Tabla 10.-** Prueba de homogeneidad de varianza

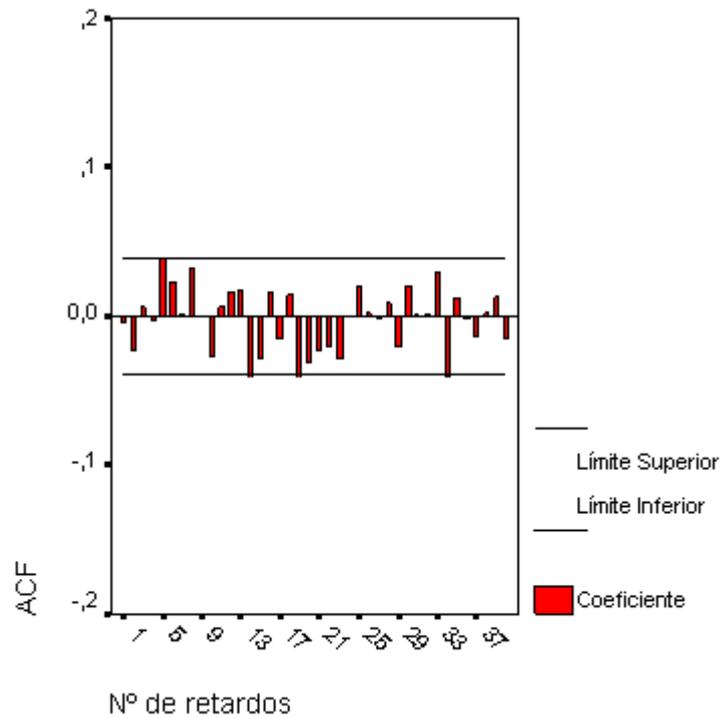
**Prueba de homogeneidad de varianzas**

Error for CAUDAL from ARIMA, MOD\_9 NOCON

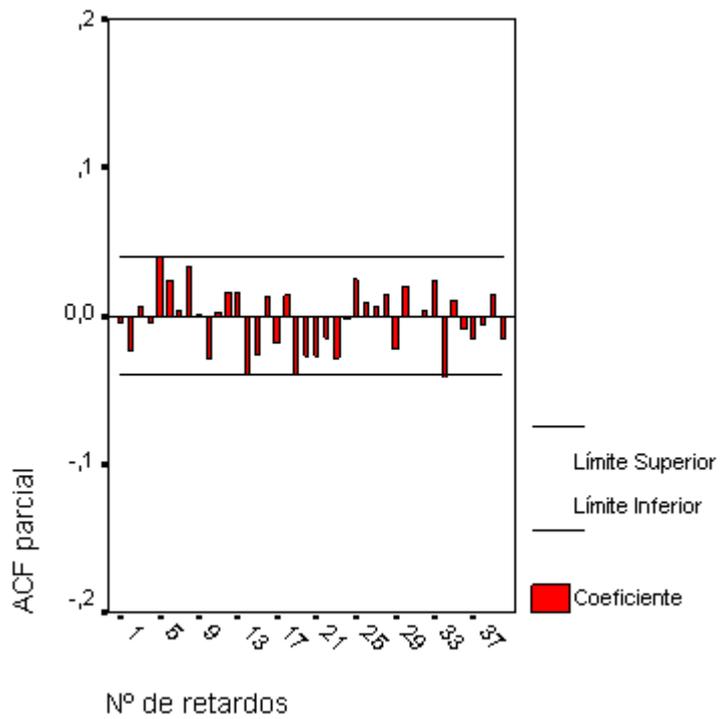
Estadístico de Levene	gl1	gl2	Sig.
2,627	6	2519	,015

También se muestra la prueba de Levene, de acuerdo a la significancia (última columna) no hay suficiente evidencia para rechazar la hipótesis nula de homogeneidad de varianzas al 0,01 nivel de significancia.

Ahora probaremos que los residuos no están correlacionados entre si.



**GRAFICO 16.-** Diagrama de autocorrelación simple de los errores o residuos



**GRAFICO 17.-** Diagrama de autocorrelación parcial de los errores o residuos

Los diagramas de autocorrelación simple y parcial no muestra autocorrelación en los residuos.

**Tabla 11.-** Prueba de Kolmogorov-Smirnov para una muestra

		ERROR
Parámetros normales <sup>a,b</sup>	Media	62,1088
	Desviación típica	96,26291
Diferencias más extremas	Absoluta	,144
	Positiva	,144
	Negativa	-,091
Z de Kolmogorov-Smirnov		,596
Sig. asintót. (bilateral)		,870

a. La distribución de contraste es la Normal.

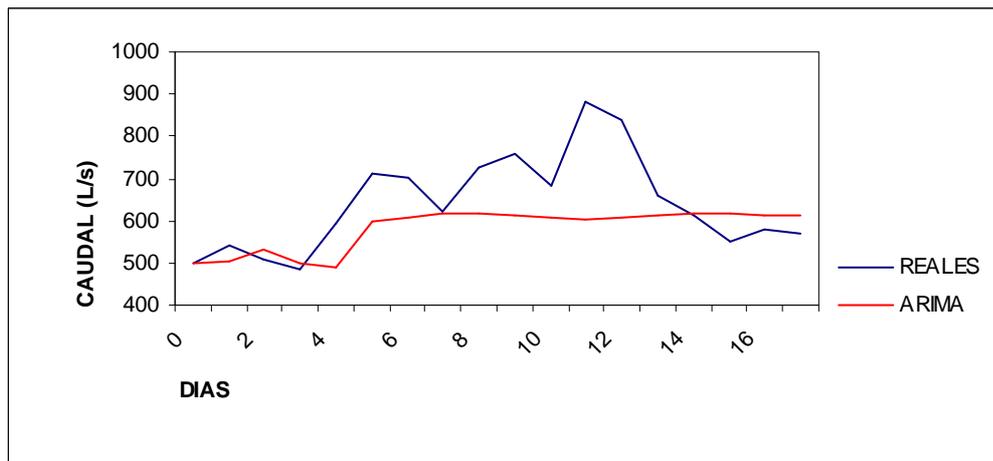
b. Se han calculado a partir de los datos.

Observe tabla 4 no existe suficiente evidencia para rechazar la hipótesis nula de que el ruido blanco sigue una distribución normal.

Se a demostrado que cumple las condiciones de ruido blanco, por ello procedemos a predecir.

### 3.1.6. PREDICCIÓN

En muchos casos las predicciones obtenidas por éste método son más contables particularmente para predicciones de corto plazo.



**GRAFICO 18.-** Datos de los caudales originales y la predicción ARIMA durante 17 días.

El modelo desarrollado permite realizar predicciones para períodos de tiempo anteriores y posteriores al considerado en este análisis (2000-2006). En el gráfico N° 18 puede observarse la excesiva variabilidad de la serie, lo que dificulta la obtención de un modelo que contemple tales variaciones.

**Tabla 12.-** Correlación de los datos reales y predichos durante 17 días.

		REAL	ARIMA
REAL	Correlación de Pearson	1	,517*
	Sig. (bilateral)	,	,034
	N	17	17
ARIMA	Correlación de Pearson	,517*	1
	Sig. (bilateral)	,034	,
	N	17	17

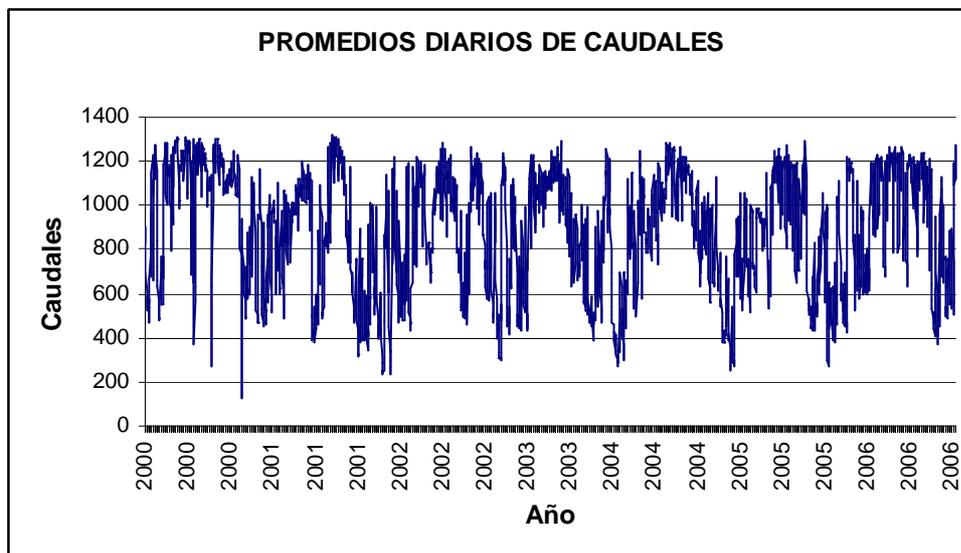
\*. La correlación es significativa al nivel 0,05 (bilateral).

La tabla 11 de correlación muestra un 0,517 de correlación entre los datos originales y los datos predichos al 0,05 nivel de significancia.

### 3.2. MODELO TISEAN 2.1

El modelo usado en el procesamiento de los datos es el TISEAN 2.1 (Noviembre 2000), que permite el análisis de series de datos no lineales, es un software de libre uso que esta disponible en internet, desarrollado por, Rainer Hegger, Holger Kantz de la Universidad de Frankfurt y Thomas Schreiber del Instituto Max-Pank.

En éste estudio básicamente se ha usados el modelo de mutua información para el calculo del tiempo de retardo, falsos vecinos para la dimensión de encaje, espectro de Lyapunov para el calculo de los coeficientes de Lyapunov y el modelo de funciones radiales para la predicción.

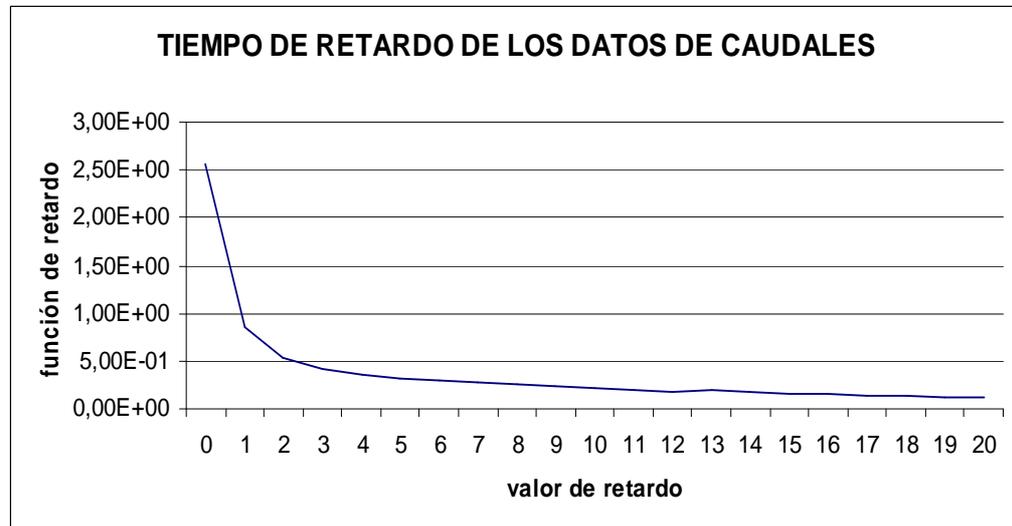


**Fuente:** Estación Nº 2 EERSA.

**GRAFICO 19.-** Promedios diarios de los caudales de río Blanco (2000-2007).

Al igual que para el análisis con la metodología de Box-Jenkins se utilizará como datos los promedios diarios de los caudales de río Blanco (2000-2007).

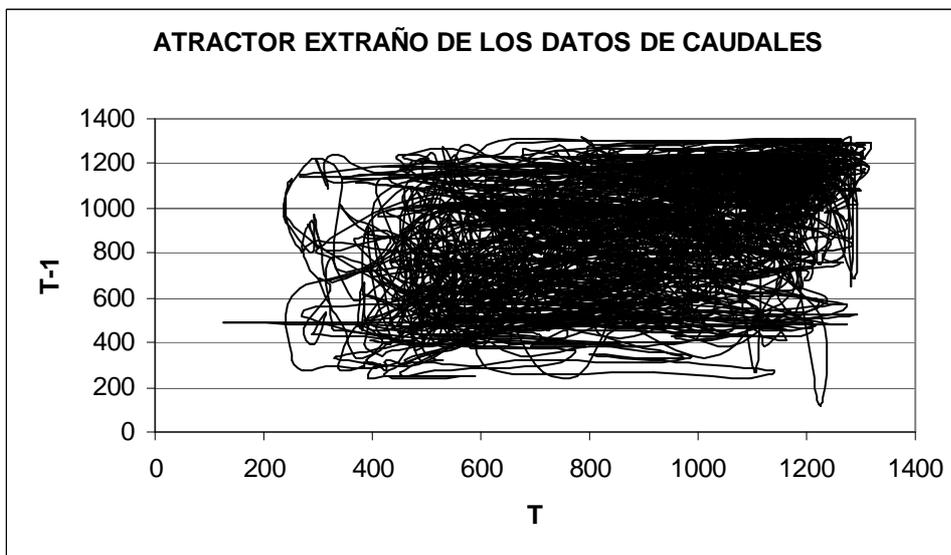
### 3.2.1. TIEMPO DE RETARDO SIN REDUCCIÓN DEL RUIDO



**GRAFICO 20.-** Tiempo de retardo de los datos de caudales sin reducción del ruido.

El primer mínimo de la función de correlación de retardo aparece en el número **12** lo que vendría a ser el tiempo de retardo.

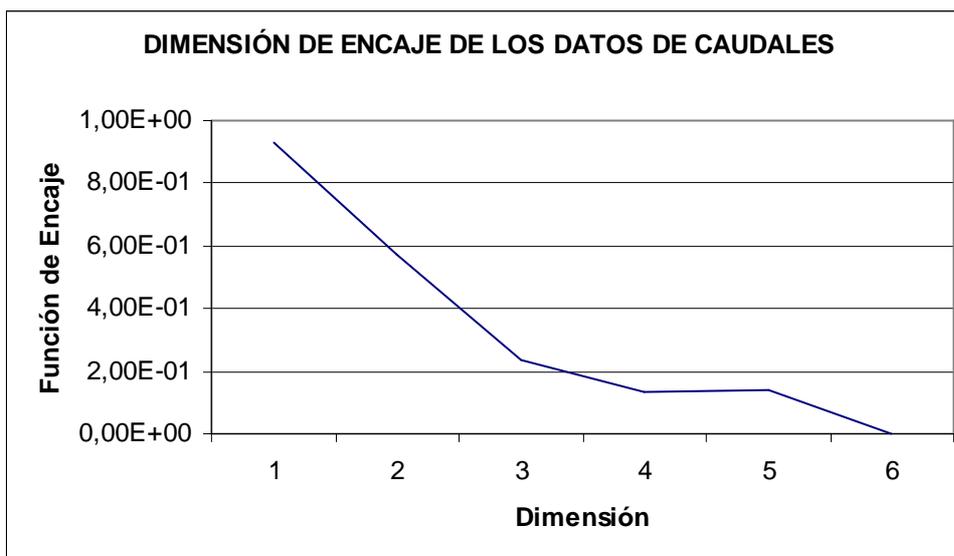
### 3.2.2. ATRACTOR EXTRAÑO



**GRAFICO 21.-** Proyección bidimensional del atractor extraño de los datos de caudal sin reducción del ruido.

Normalmente un atractor muestra cierta simetría pero en el atractor extraño de los promedios diarios de los caudales no se observa tendencia por lo que es necesario reducir el ruido.

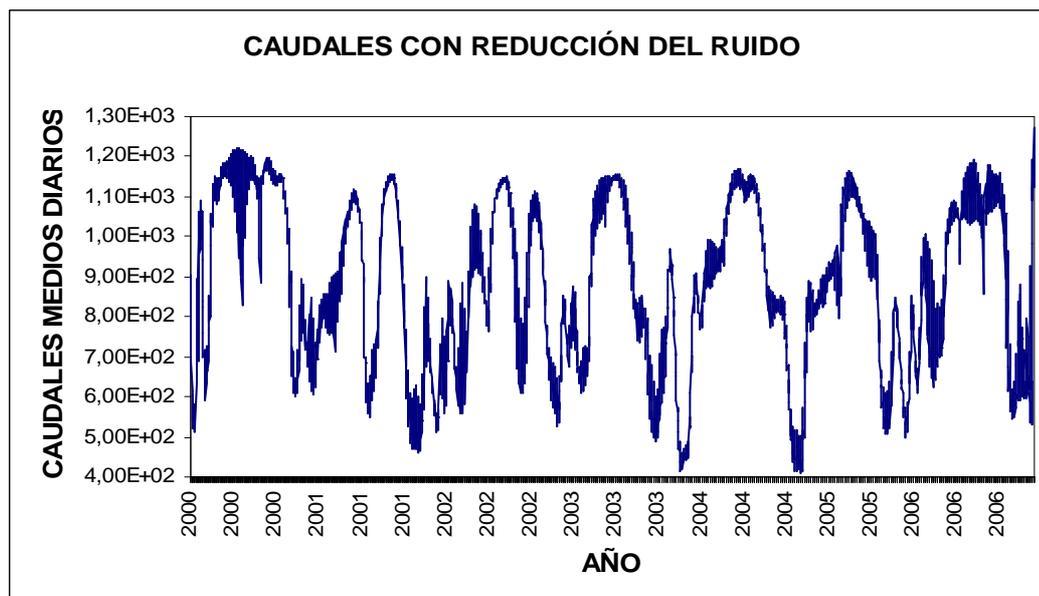
### 3.2.3. DIMENSIÓN DE ENCAJE (EMBENDING)



**GRAFICO 22.-** Gráfico de la dimensión de encaje de los datos de caudal sin reducción del ruido.

La dimensión de encaje para los datos de caudales es 6, es decir, la dimensión mínima en la que se le puede representar los datos es 6.

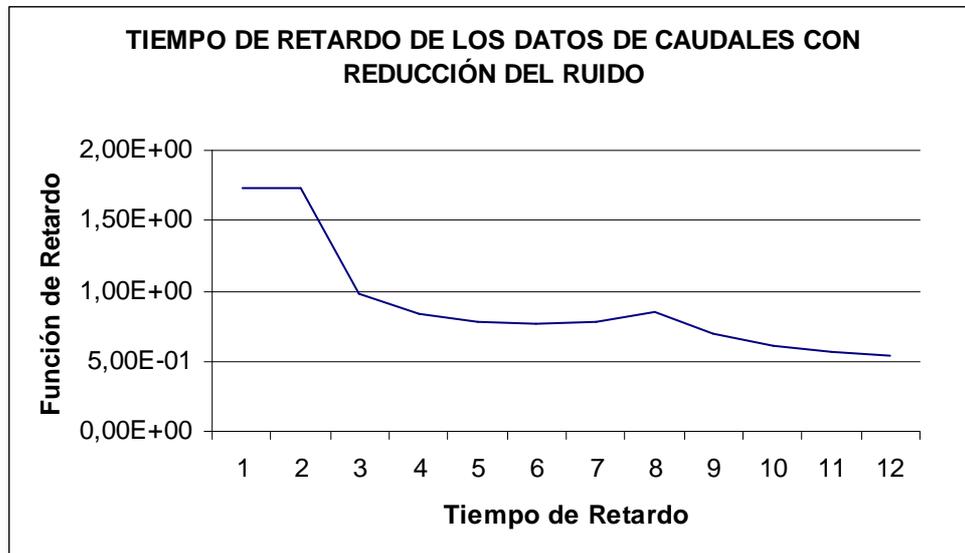
### 3.2.4. DATOS CON REDUCCIÓN DE RUIDO



**GRAFICO 23.-** Gráfico de los datos de caudal con reducción del ruido.

Una vez reducido el ruido se puede observar en la gráfica que los datos se estabilizaron.

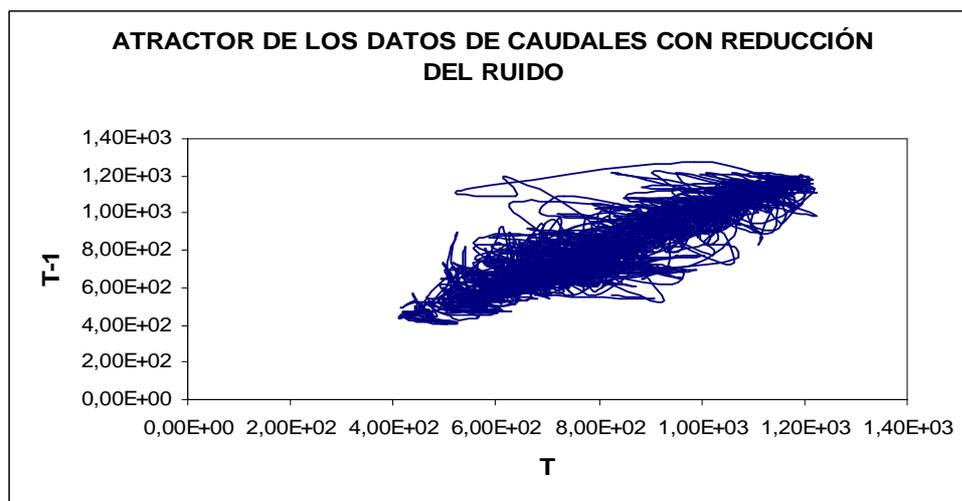
### 3.2.5. TIEMPO DE RETARDO DE SERIE DE DATOS CON REDUCCIÓN DEL RUIDO



**GRAFICO 24:** Gráfico de la función de retardo (mutual) que determina el tiempo de retardo para la reconstrucción del espacio de las fases de los datos de caudal con reducción del ruido.

El tiempo de retardo determinado de la función de retardo para los siguientes parámetros es 4.

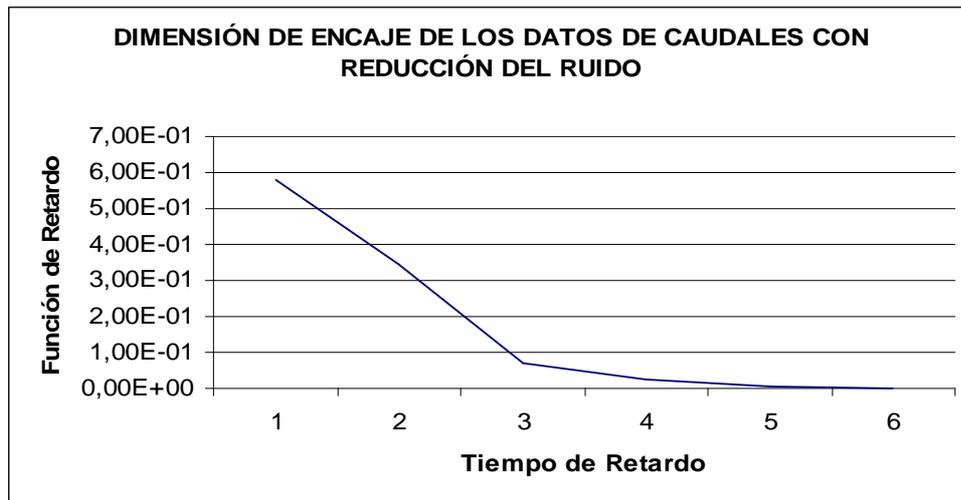
### 3.2.6. ATRACTOR EXTRAÑO CON REDUCCIÓN DEL RUIDO



**GRAFICO 25.-** Proyección bidimensional del atractor extraño de los datos de caudales con reducción del ruido.

Una vez reducido el ruido, el gráfico del atractor presenta un comportamiento mas regular que el atractor sin reducción del ruido.

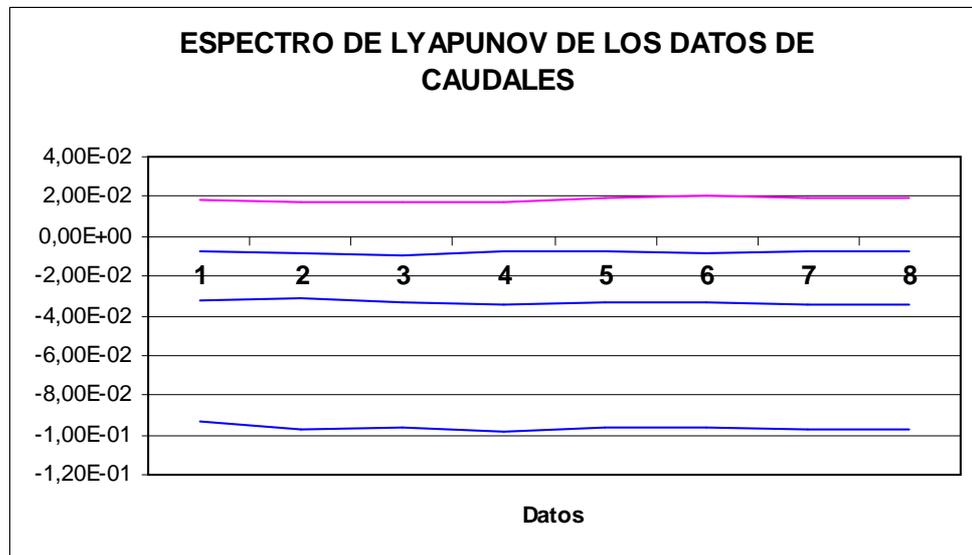
### 3.2.7. DIMENSION DE ENCAJE CON REDUCCIÓN DEL RUIDO



**GRAFICO 26.-** Gráfico de la función de encaje (falsos vecinos) que determina la dimensión de encaje para la reconstrucción del espacio de las fases de los datos de caudales con reducción del ruido.

La dimensión de encaje determinada con reducción del ruido para los caudales es 6.

### 3.2.8. ESPECTRO DE LYAPUNOV

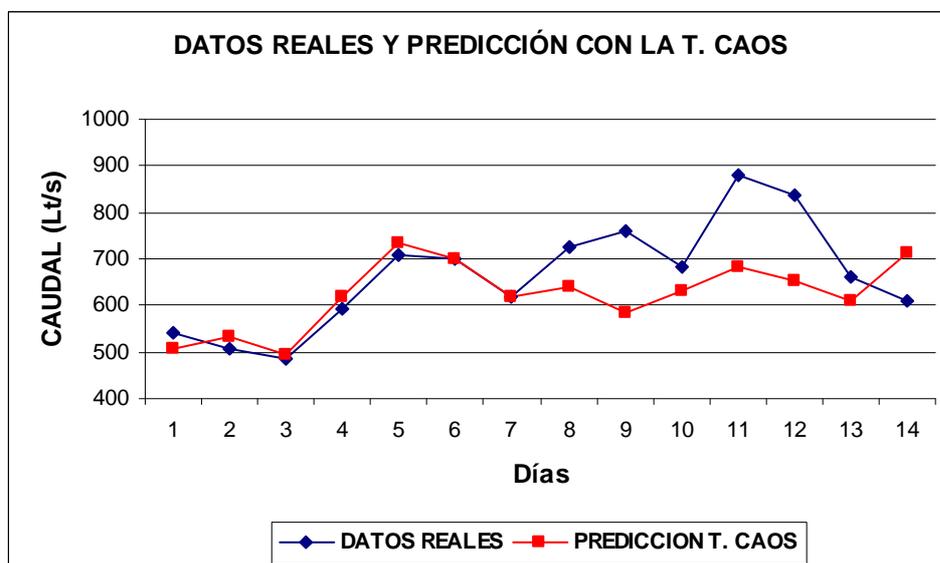


**GRAFICO 27.-** Gráfico del espectro de Lyapunov de los datos de caudales con reducción del ruido.

El gráfico del espectro de Lyapunov de los datos de caudales con reducción del ruido muestra un coeficiente positivo por lo que concluimos que la serie de los caudales es caótica.

El valor de la dimensión fractal es 2,331 y la entropía del sistema es 0,02.

### 3.2.9. PREDICCIÓN



**GRAFICO 28.-** Datos reales y predichos de los caudales de río Blanco durante los 14 días

En el gráfico 28 observe que los datos predichos coinciden solamente en dos puntos, hacemos el análisis de correlación para conocer el parentesco entre los datos reales y lo predichos.

**Tabla 13.-** Correlación de los datos reales y predichos.

		REAL	T.CAOS
REAL	Correlación de Pearson	1	,631*
	Sig. (bilateral)	,	,016
	N	14	14
T.CAOS	Correlación de Pearson	,631*	1
	Sig. (bilateral)	,016	,
	N	14	14

\*. La correlación es significativa al nivel 0,05 (bilateral).

La correlación existente entre los datos originales y los predichos, es decir mientras más largo sea el período de predicción menos se ajustarán los datos predichos a los datos reales.

**CAPÍTULO IV**

**ANÁLISIS DE**

**PRECISIÓN**

#### 4. ANÁLISIS DE PRECISIÓN

Se hace un análisis de comparación de media de los datos predichos durante los 17 días para saber si en este período de tiempo se puede apreciar la precisión de los modelos.

**Tabla 14.-** Prueba de Tukey para comparación de medias durante 17 días

HSD de Tukey<sup>d</sup>

MODELO	N	Subconjunto para alfa = ,05
		1
ARIMA	17	586,2441
T. CAOS	17	638,1964
REAL	17	648,3529
Sig.		,082

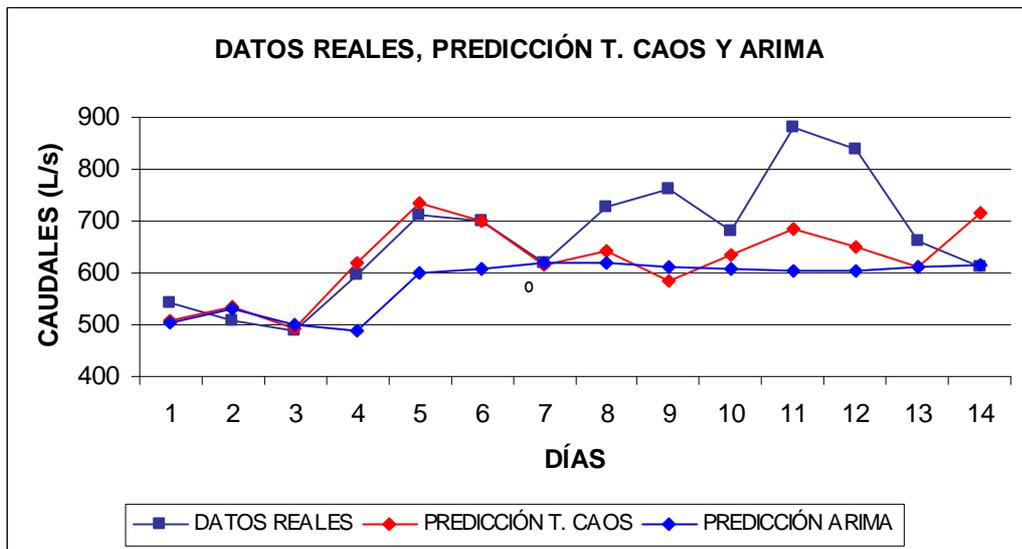
Se muestran las medias para los grupos en los subconjuntos homogéneos.

Al 0,05 nivel de significancia no hay evidencia suficiente para rechazar la hipótesis nula, es decir las medias de los datos originales y los datos predichos en un período de 17 días con el modelo ARIMA y la teoría del caos son estadísticamente iguales.

Para apreciar de mejor manera la precisión de los pronósticos con los dos métodos utilizados se realizará el análisis minucioso de los pronósticos para apreciar de mejor manera su evolución.

## 4.1. COMPARACIÓN DE LOS PRONÓSTICOS OBTENIDOS CON EL MODELOS ARIMA Y LA TEORÍA DEL CAOS DURANTE LOS PRIMEROS 14 DÍAS

### 4.1.1. GRÁFICO DE LAS SERIES



**GRAFICO 29.-** Gráfico de los datos reales & predicción ARIMA & Teoría del Caos en los primeros 14 días

A simple vista el gráfico 29 muestra que la línea de pronóstico con la teoría del caos está casi sobre la línea de los datos reales, podría decirse que a simple vista los pronósticos con la teoría del caos es más precisa que los pronósticos con el modelo, más se adelante hace un análisis para comprobarlo.

#### 4.1.2. ANÁLISIS DE CORRELACIÓN

**Tabla 15.-** Correlación de los datos reales y predichos.

		REAL
REAL	Correlación de Pearson	1
	Sig. (bilateral)	.
	N	14
ARIMA	Correlación de Pearson	,685**
	Sig. (bilateral)	,007
	N	14
T.CAOS	Correlación de Pearson	,631*
	Sig. (bilateral)	,016
	N	14

\*\* . La correlación es significativa al nivel 0,01 (bilateral).

\* . La correlación es significativa al nivel 0,05 (bilateral).

La correlación de las predicciones con el modelo ARIMA es mayor que la correlación de las predicciones con la teoría del caos respecto a los datos originales.

#### 4.1.3. ANALISIS DE LOS RESIDUOS O ERRORES

1.- Error porcentual medio	ARIMA	7,88%
	T. CAOS	3,52%
2.- Error medio cuadrático	ARIMA	0,016
	T. CAOS	0,010
3.- Error medio absoluto	ARIMA	63,85
	T. CAOS	48,32
4.- Error absoluto porcentual medio	ARIMA	8,65%
	T. CAOS	6,60%

Los errores de las predicciones con la teoría del caos durante los primeros 14 días son menores a los del modelo ARIMA, es necesario hacer una prueba de Tukey para saber si la diferencia entre los errores es significativa.

#### 4.1.4. COMPARACIÓN DE MEDIAS DE LAS PREDICCIONES

**Tabla 16.-** Prueba de Tukey para comparación de medias durante 14 días de pronóstico

HSD de Tukey<sup>a</sup>

MODELO	N	Subconjunto para alfa = ,05	
		1	2
ARIMA	14	580,288	
T. CAOS	14	622,988	622,988
REAL	14		665,714
Sig.		,385	,384

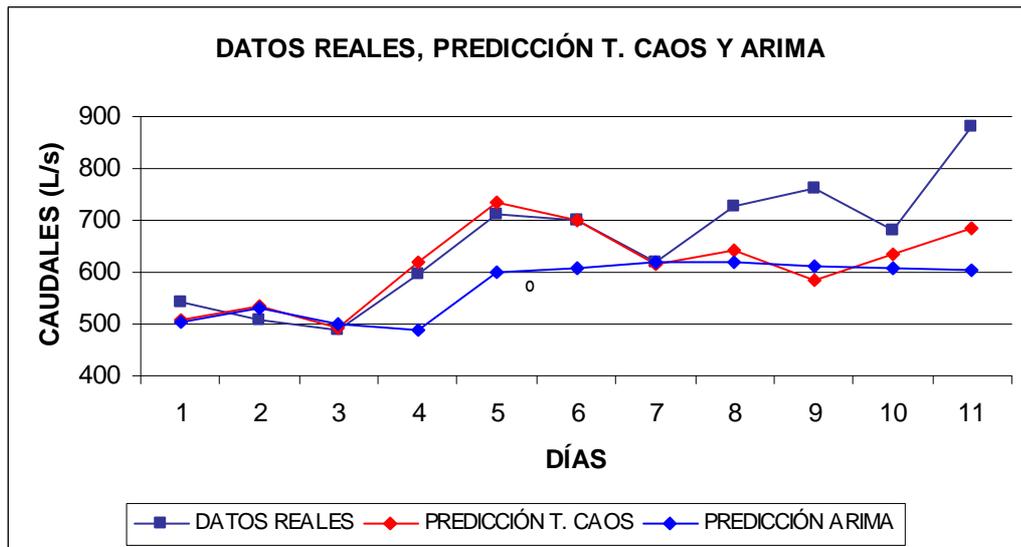
Se muestran las medias para los grupos en los subconjuntos homogéneos.

La tabla 14 muestra la prueba de Tukey misma que al 95% de confiabilidad existe diferencia significativa entre las medias de los pronósticos con el modelo ARIMA y los datos reales, mientras que son estadísticamente iguales los pronósticos con la teoría del caos y los datos reales.

Se realiza un análisis sobre los primeros 11 días de pronóstico para saber si durante éste período también muestra diferencias significativas entre los pronósticos.

## 4.2. COMPARACIÓN DE LOS PRONÓSTICOS OBTENIDOS CON EL MODELOS ARIMA Y LA TEORÍA DEL CAOS DURANTE LOS PRIMEROS 11 DÍAS

### 4.2.1. GRÁFICO DE LAS SERIES



**GRAFICO 30.-** Gráfico de los datos reales & predicción ARIMA & Teoría del caos en los primeros 11 días

No te que a partir del cuarto día el pronóstico con la teoría del caos sigue la misma trayectoria que los datos reales asta el séptimo día, se hará un análisis de correlación y una prueba de comparación de las medias de los datos predichos y los datos originales.

#### 4.2.2. ANÁLISIS DE CORRELACIÓN

**Tabla 17.-** Correlación de los datos reales y predichos.

		REAL
REAL	Correlación de Pearson	1
	Sig. (bilateral)	,
	N	11
ARIMA	Correlación de Pearson	,761**
	Sig. (bilateral)	,007
	N	11
T.CAOS	Correlación de Pearson	,746**
	Sig. (bilateral)	,008
	N	11

\*\* . La correlación es significativa al nivel 0,01 (bilateral).

\* . La correlación es significante al nivel 0,05 (bilateral).

La tabla 16 muestra que la correlación de los datos predichos con la teoría de caos a un período de tiempo de 11 días es 0,746 y con el modelo ARIMA es 0,761.

#### 4.2.3. ANALISIS DE LOS RESIDUOS O ERRORES

1.- Error porcentual medio	ARIMA	6,16%
	T. CAOS	2,87%
2.- Error medio cuadrático	ARIMA	0,012
	T. CAOS	0,006
3.- Error medio absoluto	ARIMA	49,57
	T. CAOS	31,39
4.- Error absoluto porcentual medio	ARIMA	6,88%
	T. CAOS	4,28%

Aún en 11 días los residuos que se tiene con la teoría del caos son menores que los obtenidos con el modelo ARIMA, además, note que mientras más corto sea el período de tiempo del pronóstico menores son los residuos o errores.

**Tabla 18.-** Prueba de Tukey para comparación de medias durante 11 días de pronóstico

HSD de Tukey<sup>a</sup>

MODELO	N	Subconjunto para alfa = ,05	
		1	
ARIMA	11	571,8101	
T. CAOS	11	613,1945	
REAL	11	655,3636	
Sig.			,082

Se muestran las medias para los grupos en los subconjuntos homogéneos.

En un período de pronóstico de 11 días no existe diferencia significativa entre los dos métodos utilizados al 95% de confiabilidad.

**OBSERVACION:**

**Tabla 19.-** Prueba de comparación de medias de pronósticos a un período de 12 días.

HSD de Tukey<sup>a</sup>

MODELO	N	Subconjunto para alfa = ,05	
		1	2
ARIMA	12	574,6209	
T. CAOS	12	616,4006	616,4006
REAL	12		670,6667
Sig.		,495	,311

Se muestran las medias para los grupos en los subconjuntos homogéneos.

Observe que son estadísticamente iguales los datos predichos con la teoría del caos y los datos reales.

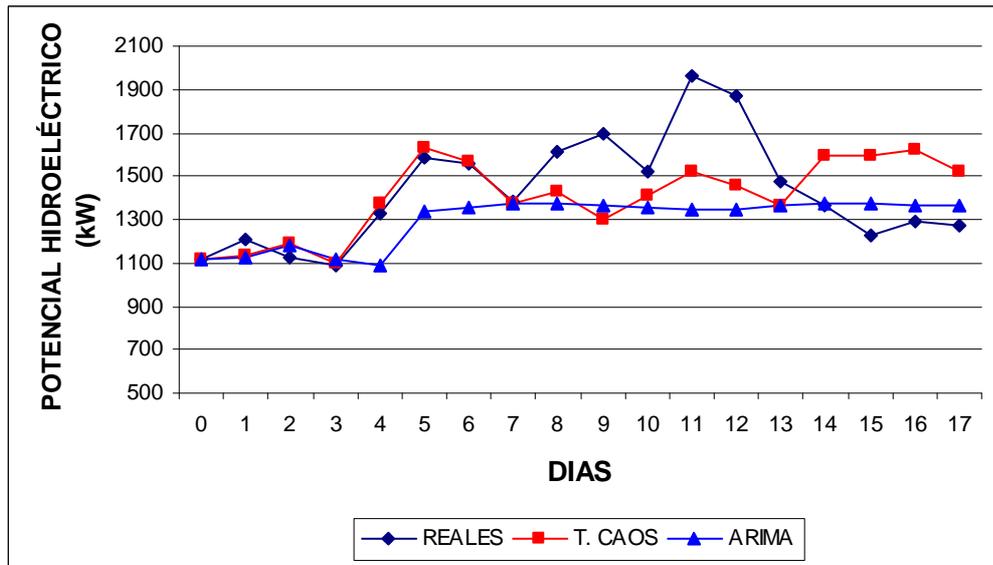
### 4.3. PRONÓSTICO DEL POTENCIAL ENERGÉTICO HÍDRICO DE RÍO BLANCO

Para hallar el potencial energético hídrico del río Blanco transformamos los caudales medios diarios de litros por segundo (L/s) a metros cúbicos por segundo ( $m^3 /s$ ).

**Tabla 20.-** Predicciones con el modelo ARIMA y teoría del caos del potencial energético hídrico

Nº	POTENCIAL HIDROELÉCTRICO (kW)		
	REALES	T. CAOS	ARIMA
1	1114,34	1114,34	1114,34
2	1208,87	1130,59	1123,81
3	1128,82	1188,23	1181,91
4	1085,64	1100,32	1112,68
5	1324,58	1376,66	1089,58
6	1581,71	1634,48	1338,24
7	1560,54	1563,79	1354,14
8	1381,32	1374,46	1376,60
9	1615,88	1429,27	1376,20
10	1696,95	1301,80	1362,02
11	1518,10	1410,93	1354,85
12	1965,32	1522,21	1348,15
13	1869,40	1452,35	1349,55
14	1471,48	1361,29	1364,32
15	1364,23	1591,76	1373,81
16	1227,54	1595,70	1376,44
17	1294,40	1624,28	1365,87
<b>MEDIA</b>	<b>1445,05</b>	<b>1422,33</b>	<b>1306,55</b>

El potencial energético hídrico del río blanco se halla con los pronósticos de los caudales en metros cúbicos, salto bruto 284,27 m y la constante  $7,84 \text{ kg /m}^2 \text{ s}^2$



**GRAFICO 31.-** Pronósticos de potencial energético hídrico del río Blanco usando la teoría del caos y el modelo ARIMA

**Tabla 21.-** Prueba de Tukey para comparación de las medias del pronóstico del potencial energético hídrico del río Blanco

HSD de Tukey<sup>a</sup>

MODELO	N	Subconjunto para alfa = ,05
		1
ARIMA	17	1306,5488
T. CAOS	17	1422,3335
REAL	17	1445,0512
Sig.		,082

Se muestran las medias para los grupos en los subconjuntos homogéneos.

Las medias de los pronósticos del potencial energético hídrico con los dos métodos y los datos reales son estadísticamente iguales al 95% de confiabilidad.

## CONCLUSIONES

- La reducción del ruido en el análisis con la teoría del caos redujo la irregularidad del atractor.
- Los datos de caudales analizados son caóticos por presentar un coeficiente positivo de Lyapunov.
- Una diferenciación en el análisis del modelo ARIMA acabó con el problema de no estacionariedad en media.
- Se logró mayor correlación, con la teoría del caos 0,97 a un período de 7 días y con el modelo ARIMA 0,83 en período de 10 días.
- Los residuos o errores obtenidos en las predicciones con la teoría del caos son menores que los obtenidos con el modelo ARIMA.
- Las predicciones son mas precisas mientras más pequeño es el período de predicción.
- El período de predicción con correlación significativa al 5% es de 14 días con la teoría del caos y 17 días con el modelo ARIMA.

- No existe diferencia significativa entre las predicciones con los dos métodos y los datos reales hasta un período de pronóstico de 11 días (tabla 17) a un nivel de significancia de 0,05.
  
- Las medias de los pronósticos del potencial energético hídrico de río Blanco durante 11 días obtenidos con los dos métodos son estadísticamente iguales al potencial hallado con los datos reales al 95% de confiabilidad.
  
- Existe diferencia significativa en las predicciones a partir del doceavo día al 95% de confiabilidad, es decir, los pronósticos con la teoría del caos es más precisa que los pronósticos hallados con el modelo ARIMA a partir del doceavo día.

## RECOMENDACIONES

- Implementar una base de datos informática para un mejor manejo de los datos en la Empresa Eléctrica Riobamba S.A.
- Fomentar el uso de modelos de pronósticos para la toma de decisiones con respecto a las diferentes actividades que se realizan en las centrales hidroeléctricas
- Realizar estudios de potencial hidroeléctrico en la provincia de Chimborazo para mediante una base de datos histórica aprovechar de mejor manera las fuentes de energía hídrica.
- Utilizar caudalímetro para reducir el error en la medición de los caudales.
- Fomentar el uso de modelos de pronósticos para estudios de potencial hidroeléctrico en la provincia de Chimborazo para su mejor aprovechamiento.
- Estudiar la teoría del caos en los métodos de pronósticos que se ilustran en la asignatura Series Temporales en la carrera de Estadística Informática.

## **RESUMEN**

El objetivo de ésta investigación fue modelar y predecir con mayor precisión el potencial energético hídrico que posee el río Blanco, por medio de la predicción de caudales en L/s en base a los registros almacenados en la estación 2 de la Empresa Eléctrica Riobamba S.A. aplicándose modelos ARIMA y teoría del caos cuyos resultados son importantes para el proyecto “Balance Energético y Diagnóstico del Potencial de Energías Alternativas del Cantón Riobamba” financiado por el CONESUP, tendiente a resolver problemas de aprovechamiento de energía hidráulica.

Se utilizó el software SPSS 11 para el modelo ARIMA y TISEAN 2.1 para teoría del caos, la precisión de pronóstico se halla mediante el análisis de residuos o errores, correlación y prueba de Tukey. El período de predicción con correlación significativa es de 14 días con la teoría del caos y 17 días con el modelo ARIMA, la prueba de Tukey muestra que no existe diferencia significativa entre las predicciones y los datos reales hasta un período de pronóstico de 11 días (tabla 17), es decir, las medias de los pronósticos del potencial energético hídrico de Río Blanco hallados con los caudales predichos con los dos métodos son estadísticamente iguales al potencial hallado con los datos reales durante 11 días al 95% de confiabilidad, a partir del doceavo día la teoría del caos es más precisa que el modelo ARIMA. La Empresa Eléctrica Riobamba S.A debería Implementar una base de datos informática para un mejor manejo de información y fomentar estudios de potencial hidroeléctrico mediante modelos de pronósticos para su mejor aprovechamiento.

## **SUMMARY**

The objective of this investigation was to model and to predict with more precision the potential hydroelectric energy of the Río Blanco, by means of the prediction of flows in L/s based on the registrations stored in station number 2 of the Empresa Eléctrica Riobamba S.A. applying the ARIMA model and chaos theory whose results are important for the project “Energy Balance and Diagnosis of the Potential of Alternative Energy of the Canton Riobamba” financed by the CONESUP, with the end of solving the problems of using of hydroelectric energy.

The software SPSS version 11 for the ARIMA model and TISEAN version 2.1 for chaos theory, the presage precision was determined by an residuals or errors analysis, correlation and Tukey test. The forecasting period with significant correlation is 14 days with the chaos theory and 17 days with the ARIMA model, the Tukey test shows that significant difference don't exist between the predictions and the real data until day 11 (chart 17), that is to say, the averages of the precision of the potential hydroelectric energy of Río Blanco found with predicted with the two methods is statistically similar to the potential found with the real data during 11 days with reliability at 95%, starting from the twelfth day the chaos theory is more precise than the model ARIMA. The Empresa Eléctrica Riobamba S.A should implement a computer database for a better handling of information and to start studies of hydroelectric potential by means predict models for the most efficient use of the river.

## BIBLIOGRAFÍA

- Cabrera J. Carlos y Jaramillo J. Jorge, Estadística Inferencial, Loja: Universidad Técnica de Loja. s.f. (Documento).
- Daza P. Jorge F. Análisis Estadístico con SPSS 14. ed. Lima: Megabyte, 2006.
- Freedman David H., Chaos Theory Inc. ed. Boston: s.e. Oct 20, 1998.
- Gujarati Damodar N. Econometría. 4ta, ed. McGraw-Hill, s.f. (Fotocopias)
- Hayles, La Evolución del Caos, El Orden Dentro del Desorden en las Ciencias Contemporáneas. ed. Barcelona: Gedisa N. K. 1993.
- Honke John E., Reitsch Arthur G., Pronósticos en los Negocios. 5ª. ed. Presentice Hall, s.f.
- Kathleen T. Alligood, Chaos an Introduction to Dynamical Systems Springer-Verlag. ed. New York, 1996.
- Mª De Juana José (coordinador), Energías Renovables para el Desarrollo, Cooperación Internacional. s.e. Madrid, s.f.
- Madrid, Cedeti, Ministerio de Industria y Energía, Pequeñas Centrales Hidroeléctricas, Centro para el Desarrollo Tecnológico Industrial. ed. Progenza, 2000, (Documento).
- Mandelbrot BB. La Geometría Fractal de la Naturaleza. ed. Barcelona: Tusquets, 1997.
- Mencuccini Silvestrini. Física II. ed. Napoli-Italia: Liguori, 1988.
- Ortega M. Energías Renovables. ed. Madrid: Paraninfo, 2002.

- Patiño JF. Caos y Complejidad, Las Ciencias del siglo XXI. Lecturas de Nutrición 6, 2000, (Fotocopias).
- Pérez López César, Técnicas Estadísticas con SPSS. Universidad Complutense de Madrid, Instituto de Estudios Fiscales: Prentice Hall, 2001
- Pindyck R. S. Econometría, Modelos y Pronósticos. ed. McGraw-Hill, 2001.
- Rasband, Chaotic Dynamics of non Linear Systems, Jhon Wiley & Sons, 1990.
- Semansky, Física II, ed. México: Cagsa, 2004.
- Sanz Badia, Energía Hidroeléctrica, ed. Zaragoza: Centro Politécnico Superior, 2002.
- Ing. J.Z., Informe del Proyecto Hidroeléctrico Quimiag Río Blanco, Tomo I (5-Q-1), Quito, s.f. (Documento).

## BIBLIOGRAFÍA EN INTERNET

➤ Disipación de energía

<http://www.unesco.org.uy/phi/libros/obrashidraul/Cap6-d.html>

11/04/2006

➤ Energía Hídrica APROTEC

[http://www.APROTEC7EnergiaHidrica\\_micri.htm](http://www.APROTEC7EnergiaHidrica_micri.htm)

15/04/2006

➤ Energía Hidroeléctrica

[http://www.consumer.es/web/es/medio\\_ambiente/energia\\_y\\_ciencia/2004/08/09/140155.php](http://www.consumer.es/web/es/medio_ambiente/energia_y_ciencia/2004/08/09/140155.php)

13/04/2006

➤ Generalidades de Pronósticos

<http://www.gestiopolis.com/recursos/documentos/fulldocs/ger1/serietiempo.htm#tipo>

25/05/2006

➤ “La Predicción Y La Teoría Del Caos”,

<http://www.aedemo.es/septiembre97/privado/caos.html>

11/04/2006

➤ Series de Tiempo ARIMA

<http://www.einsteinnet.com/econometria/ARIMA/arima1.htm>

25/05/2006

➤ Modelos ARIMA

[http://www.uam.es/docencia/predysim/combinado4/4\\_2\\_ficha.htm](http://www.uam.es/docencia/predysim/combinado4/4_2_ficha.htm)

04/09/2006

➤ Modelos Box-Jenkins

<http://www.usuarios.lycos.es/guillemat/box-jenkins.htm>

<http://www.ucm.es/info/ecocuan/mjm/ecoaplimj/BoxJenkins.pdf> (pdf)

5/09/2006

➤ Técnicas de Pronósticos

<http://www.soyentrepreneur.com/pagina.hts?N=14912>

5/09/2006

# **ANEXOS**

**ANEXO 1:** Caudales reales y predichos, modelo ARIMA y Teoría del CAOS durante 17 días

Nº	DATOS REALES	PREDICCIÓN (L/s)	
		T. CAOS	ARIMA
0	500	500,000	500,000
1	542	507,292	504,250
2	507	533,157	530,317
3	487	493,711	499,256
4	594	617,702	488,893
5	710	733,384	600,464
6	700	701,667	607,596
7	620	616,717	617,677
8	725	641,308	617,498
9	761	584,113	611,134
10	681	633,079	607,916
11	882	683,010	604,910
12	839	651,667	605,540
13	660	610,808	612,165
14	612	714,218	616,422
15	551	715,984	617,605
16	581	728,807	612,862
17	570	682,714	611,645
<b>MEDIA</b>	<b>648,39</b>	<b>638,20</b>	<b>586,24</b>

**ANEXO 2:** Central Hidroeléctrica río Blanco Quimiag – Riobamba



**Fig. 1:** Central Hidroeléctrica río Blanco Quimiag – Riobamba



**Fig. 2:** Tableros de control central río Blanco Quimiag



**Fig. 3:** Turbina central río Blanco Quimiag



**Fig. 4:** Cuarto de control central río Blanco



**Fig. 5:** Tubería de presión central río Blanco



**Fig. 6:** Desarenador canal de riego río Blanco



**Fig. 7:** Vista general canal riego río Blanco



**Fig. 8:** Vista general tanque presión río Blanco