



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS  
PROFESIONALES "ACATLÁN"

MODELOS EN EL AJUSTE DE GASTOS MEDIOS  
DIARIOS (Qm) DE UNA ESTACIÓN HIDROMÉTRICA:  
LA ESTACIÓN TOTOLICA (1987-1994)



T E S I S

QUE PARA OBTENER EL TÍTULO DE

LIC. EN MATEMATICAS APLICADAS Y COMPUTACIÓN

P R E S E N T A

MIRIAM VELASCO VILLARREAL

*[Firma manuscrita]*

ASESOR: LIC. JUDITH JARAMILLO LOPEZ



AGOSTO DEL 2000.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## AGRADECIMIENTOS

A Dios, por permitirme vivir.

A mi madre por su infinita paciencia, comprensión y confianza en mí. A mi hermana por su respaldo incondicional y a mi familia por su apoyo.

A la Universidad Nacional por la oportunidad recibida.

A la Lic. Judith Jaramillo López por su apoyo, guía y amistad en este trabajo.

Al Lic. Jaime Ramírez Muñoz por su ayuda, respaldo y comentarios que hicieron posible la realización de este trabajo, además de su amistad y paciencia para conmigo.

A mis amigos, por su aliento y no menciono sus nombres para no omitir ninguno.

A mis maestros. Por la formación que nos dan a todos los estudiantes.

A las nuevas personas que he conocido y que sin conocerme me han prestado su ayuda como al Ing. Chávez, al Ing. Barrera Almazán del Meteorológico y la Gerencia de Aguas Superficiales respectivamente; a los compañeros de Jaime y al Colegio de México, por el préstamo de sus instalaciones. Y a todas las personas que de alguna manera han contribuido a la realización de esta obra.

## I N D I C E

	Pág.
<b>AGRADECIMIENTOS</b> .....	i
<b>INTRODUCCIÓN</b> .....	1
<b>I. DESCRIPCIÓN BÁSICA DE REDES HIDROLÓGICAS</b> .....	4
1.1 Antecedentes históricos.....	4
1.2 Descripción básica.....	8
1.2.1 Definición y objeto de la red hidrológica.....	8
1.2.2 Densidad y configuración de las redes.....	10
1.2.3 Estaciones hidrométricas.....	11
1.2.3.1 Principales.....	11
1.2.3.2 Secundarias.....	11
1.2.3.3 Para fines especiales.....	11
1.3 Frecuencia y horario de las observaciones.....	12
1.4 Cuencas.....	12
1.4.1 Definición de cuenca.....	12
1.4.2 Tipos de cuenca.....	13
<b>II. METODOS ESTADISTICOS EN ANALISIS Y PREDICCIÓN HIDROLOGICA</b> .....	14
2.1 Modelado de sistemas hidrológicos.....	14
2.2 Simulación estocástica de series hidrológicas de tiempo.....	17
2.2.1 Modelos markovianos lag-1.....	18
2.2.2 Modelos autorregresivos de promedios móviles integrados ARIMA.....	19
2.2.3 Modelos fraccionales gaussianos de ruido y de proceso de línea quebrada.....	20
2.3 Predicción hidrológica.....	20
2.3.1 Variables usuales en una predicción.....	21
2.3.2 Componentes de una predicción.....	22
2.3.3 técnicas usadas en predicción.....	22
2.3.3.1 Correlación y regresión.....	23
2.3.3.2 Análisis de series de tiempo.....	24
2.3.3.3 Técnicas de ajuste de la predicción.....	25
2.3.3.4 Predicción probabilística.....	26
2.3.3.5 Desarrollo del procedimiento de predicción.....	26
2.3.3.6 Determinación de parámetros y coeficientes del modelo.....	27
2.3.3.7 Criterios de verificación.....	28

	Pág.
<b>III. TECNICAS ESTADISTICAS.....</b>	<b>32</b>
3.1 Regresión lineal.....	32
3.1.1 Coeficiente de correlación simple.....	32
3.1.2 Regresión lineal simple.....	33
3.1.2.1 Estimación de los parámetros.....	34
3.1.2.2 Análisis de los residuos.....	34
3.1.2.2.1 Prueba Durbin-Watson.....	35
3.1.2.3 Análisis de la varianza.....	36
3.1.3 Regresión lineal múltiple.....	37
3.1.3.1 El coeficiente de correlación múltiple.....	37
3.1.3.2 Estadísticos para las variables independientes.....	39
3.2 Series temporales.....	41
3.2.1 Conceptos teóricos.....	41
3.2.1.1 Procesos estacionarios.....	41
3.2.1.2 Función de autocorrelación.....	42
3.2.1.3 Proceso estacionario.....	42
3.2.1.4 Autocorrelación simple de orden $K$ , $\phi_k$ .....	42
3.2.1.5 Autocorrelación parcial de orden $k$ , $\alpha_k$ .....	42
3.2.1.6 Proceso de ruido blanco.....	43
3.2.1.7 Proceso AR(p).....	43
3.2.1.8 Proceso MA(q).....	43
3.2.1.9 Proceso ARMA(p,q).....	44
3.2.1.10 Proceso ARIMA(p,d,q).....	45
3.2.1.11 Proceso estacional.....	45
3.2.1.11.1 Proceso ARIMA(p,d,q) (P,D,Q).....	46
3.2.2 Análisis de una serie temporal mediante la metodología Box-Jenkins.....	47
3.2.2.1 Identificación del modelo ARIMA.....	47
3.2.2.2 Homocedasticidad.....	48
3.2.2.3 Análisis de la estabilidad en medias y de la estacionalidad.....	49
3.2.2.4 Determinación de los órdenes p,d,q,P,D,Q en el modelo.....	50
3.2.2.5 Ajuste del modelo ARIMA.....	52
3.2.2.5.1 Estimación de los parámetros.....	52
3.2.2.5.2 Sobreajuste del modelo.....	53
3.2.2.5.3 Validación del modelo.....	54
3.2.2.5.4 Etapas en la construcción de un modelo ARIMA.....	54
<b>IV. LA CUENCA DEL VALLE DE MEXICO.....</b>	<b>56</b>
4.1 Breve reseña histórica.....	56
4.1.1 La época prehispánica.....	56
4.1.2 La colonia.....	58
4.1.3 En el México independiente.....	58
4.2 Conformación actual de la cuenca del valle de México.....	60
4.3 Sistema hidrológico de la cuenca del valle de México.....	62
4.3.1 Ríos del oriente.....	62
4.3.2 Ríos del sur.....	62
4.3.3 En la parte poniente.....	63

4.4 Estación hidrométrica Totolica.....	Pág. 65
4.4.1 Objetivos de su instalación.....	65
4.4.2 Características geográficas.....	66
4.4.3 Datos de instrumentación hidrométrica.....	66
4.4.4 Metodología empleada en el cálculo hidrométrico por la Gerencia de aguas del Valle de México.....	67
<b>V. GENERACION DE MODELOS EN LOS DATOS DE GASTOS MEDIOS DIARIOS DE LA ESTACION HIDROMETRICA TOTOLICA.....</b>	<b>68</b>
5.1 Introducción.....	68
5.2 Datos de entrada.....	68
5.3 Análisis por regresión lineal simple.....	69
5.4 Análisis por series de tiempo.....	72
<b>CONCLUSION.....</b>	<b>79</b>
ANEXO 1. Gráficas.....	81
ANEXO 2. Conceptos de hidrodinámica.....	89
ANEXO 3. Glosario.....	93
REFERENCIA BIBLIOGRAFICA.....	96

## INTRODUCCIÓN

Desde tiempos muy antiguos el hombre se ha preocupado por los fenómenos naturales, como huracanes, fuertes nevadas, lluvias torrenciales, sequías, etcétera, generados algunos de estos por el ciclo del agua o fenómenos atmosféricos, también por el abastecimiento de comida, agua y recursos necesarios para su supervivencia. Para esto tuvo que aprender a observar las estaciones del año, los ciclos lunares, entre otros.

En la época de los egipcios, después los romanos y otras civilizaciones tenían conocimientos del aprovechamiento de este recurso necesario para el hombre: el agua. En nuestro país los antiguos pobladores y principalmente los aztecas usaban el agua como medio de transporte para llevar y traer materiales, como medio de comunicación, para la alimentación existían las chinampas, tenían mucha relación con el agua, de ahí, que construyeron diques para separar las aguas dulces de las saladas, para controlar las inundaciones, ya que Tenochtitlán estaba asentada en una cuenca cerrada, es decir, sin salida al mar. Desde la llegada de los españoles comenzó la desecación de los grandes lagos, afectando el clima y por lo tanto la vegetación y la fauna. Cada vez se hizo más difícil traer el agua hacia la ciudad, donde se construyeron acueductos para traer el vital líquido; existieron varios proyectos para tratar de traer y desalojar al mismo tiempo el agua. Por el interés en el uso y desarrollo de tecnología sobre éste recurso natural se escogió el tema relacionado con la hidrología y al mismo tiempo aplicar conocimientos de la licenciatura.

La hidrología estudia las aguas de la superficie terrestre como la de ríos o lagos. Para medir de alguna manera el agua que fluye en los ríos, presas, etc. se han puesto en operación redes hidrológicas, conformadas por estaciones hidrométricas que proporcionan los datos para la

generación de información necesaria para la elaboración de boletines, revistas, libros, avisos, que ayuden al desarrollo y gestión de los recursos hídricos a una escala comparable con el nivel económico del país. Un ejemplo de esto puede ser la predicción, ya que la necesidad de predecir es necesaria para el aprovechamiento de los recursos de una región, como por ejemplo: evitar incendios en los bosques, abastecimiento de agua para uso doméstico, industrial e irrigación, inundaciones, etc. Los errores en predicción se pueden deber a la medición, al modelo utilizado o a los datos que varían rápidamente.

El problema planteado fue ¿Cuál de las técnicas estadísticas: Regresión lineal o Series de tiempo se ajustará a los datos hidrológicos?, los cuales se obtuvieron de la estación Totolica que se ubica a 1.5 km. al poniente del centro de la población de San Bartolo, Naucalpan, estado de México, zona densamente poblada y de gran desarrollo industrial, que necesita agua y por lo tanto es indispensable el aforo de las salidas de agua de la presa Totolica, para el aprovechamiento óptimo del recurso.

Para este trabajo se utilizaron la Regresión Lineal y Series de tiempo, el objetivo es el de comparar el ajuste de los gastos medios diarios obtenidos por medio de estas dos técnicas y lo más importante no es cuál de las dos es la mejor, sino cual es la que mejor se ajusta al conjunto de datos recolectados al confrontar las pruebas que cada una nos permite realizar.

La hipótesis principal es: Si el gasto medio diario tiene un comportamiento histórico, entonces la técnica de regresión lineal y las series temporales podrán ajustarse y explicar el comportamiento del gasto a través del tiempo proporcionado por un modelo; considerando a priori que la serie de tiempo se ajustará mejor al gasto medio diario.

La investigación esta compuesta de cinco capítulos y tres anexos. En el primer capítulo se muestra una pequeña sección de antecedentes históricos, se definen conceptos como el de redes hidrológicas, cuencas y estaciones hidrométricas, el segundo capítulo muestra algunos de los métodos estadísticos usados en análisis y predicción hidrológica; el tercer capítulo presenta la base teórica de las técnicas a aplicarse (regresión lineal y series de tiempo); el cuarto contiene una breve



reseña histórica de la cuenca del Valle de México, la conformación actual de la misma, así como la descripción de la estación Totolica, en el capítulo cinco se ponen a prueba las técnicas estadísticas citadas en el capítulo tres. En los anexos, se muestran las gráficas de las funciones de autocorrelación de los modelos utilizados, también se proporcionan conceptos de hidrodinámica y se expone un glosario de términos hidrológicos para que el lector se familiarize con los mismos.

## CAPÍTULO I

### DESCRIPCIÓN BÁSICA DE REDES HIDROLÓGICAS

#### 1.1 ANTECEDENTES HISTÓRICOS

Desde el principio de los tiempos el hombre se ha preocupado por la distribución y conservación óptima de los recursos naturales, siendo uno de los más importantes el agua. Es importante contar con un controlado del manejo de ella, para tener un adecuado aprovechamiento de la misma, que por fenómenos naturales también puede causar catástrofes como es el caso de las inundaciones. Tal vez el sistema más antiguo de aviso sobre una inundación fue utilizado en el antiguo Egipto. Los ingenieros egipcios observaban las lecturas de los niveles del agua en un dispositivo llamado nilómetro<sup>1</sup>. Si las lecturas mostraban niveles demasiado elevados de agua, enviaban un equipo de remeros para advertir a los residentes de las poblaciones en peligro a que desocuparan su lugar de residencia.

La medición del flujo de corrientes de agua constituye uno de los cálculos fundamentales en la administración de los recursos hidráulicos, se cree que fue Herón de Alejandría<sup>2</sup> el primero en proponer el concepto de que el gasto es igual a la velocidad multiplicada por el área de la sección transversal; éste concepto fue olvidado, dieciséis siglos después Benedetto Castelli<sup>3</sup>, el fundador de la hidrología italiana, llegó al mismo conocimiento.

Es probable que el abastecimiento de agua más antiguo (que se halla todavía en operación) fue la red de acueductos romanos. El ingeniero a

---

<sup>1</sup> Ver figura 1-1

<sup>2</sup> Herón de Alejandría. Matemático y Físico griego (Alejandría s.II); se dedicó al estudio de la mecánica, se le atribuye la invención de un reloj de arena, de la turbina que lleva su nombre y de una catapulta de aire comprimido. Descubrió la fórmula del área del triángulo en función de su perímetro. Fuente: *Diccionario Enciclopédico Universal*. Plaza&Janes Editores, 4 Vols.; 1992, Barcelona, España, Vol.3.

<sup>3</sup> Italia ss.XVI y XVII (Brescia, 1577- Roma, 1644) Físico y Benedictino italiano, alumno de Galileo en Padua (1604), fue nombrado por recomendación de éste, profesor de matemáticas en la Universidad de Pisa (1613). Enseñó matemáticas en la Universidad de Roma y publicó "Della misura delle acque correnti" (1628), considerado como el primer tratado moderno de hidráulica. En la correspondencia mantenida con Galileo relató sus célebres experiencias sobre absorción y transmisión de calor de los cuerpos blancos y negros. Fuente: *Salvat Universal Diccionario Enciclopédico*, Salvat editores, S.A., 23 Vols., 1991, Barcelona, España, Vol.5.

quien se debe esta obra fue Apio Claudio Craso<sup>4</sup>, quien la terminó en el año 312 a.C. De esta manera Roma tuvo el primer sistema municipal de alcantarillado, aunque sólo para el agua de lluvia; los desechos de Roma no se virtieron en el sistema hasta varios siglos después.

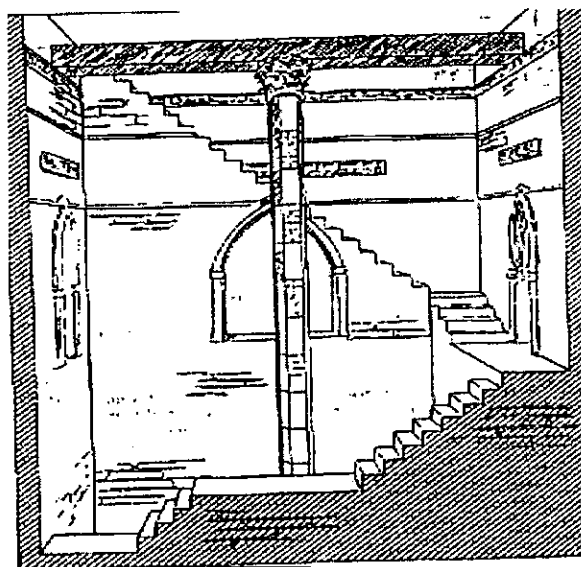


Figura 1-1 El nilómetro de Roda en 1798 d.C.

(Fuente: Helweg, Otto J. *Recursos hidráulicos: Planeación y administración*. Edit. Limusa, . Pág.14). La columna central es el nilómetro, la cual muestra líneas horizontales que son las que marcaban el nivel del agua.

Durante los siglos XVII y XVIII se formaron en Europa varios grupos: La Real Sociedad Inglesa, La Real Academia Francesa de Ciencias y el Cuerpo de Caminos y Puentes que estaban interesados en el avance de la ciencia. Uno de sus principales objetivos era tener una sólida base científica para la planeación de los recursos hidráulicos. De la última institución mencionada, hubo varios hidrólogos precursores del estudio de los recursos hidráulicos, uno de ellos fue Chézy<sup>5</sup>.

<sup>4</sup> Claudio Ceco Apio. Censor romano a.C. ocupó diversos cargos políticos, destacándose por su enemistad hacia la plebe. Mandó construir la Via Apia y el primer acueducto de Roma. Fuente: *Diccionario Enciclopédico Universal*, Plaza&Janes Editores, 4 Vols., 1992, Barcelona, España, Vol.1.

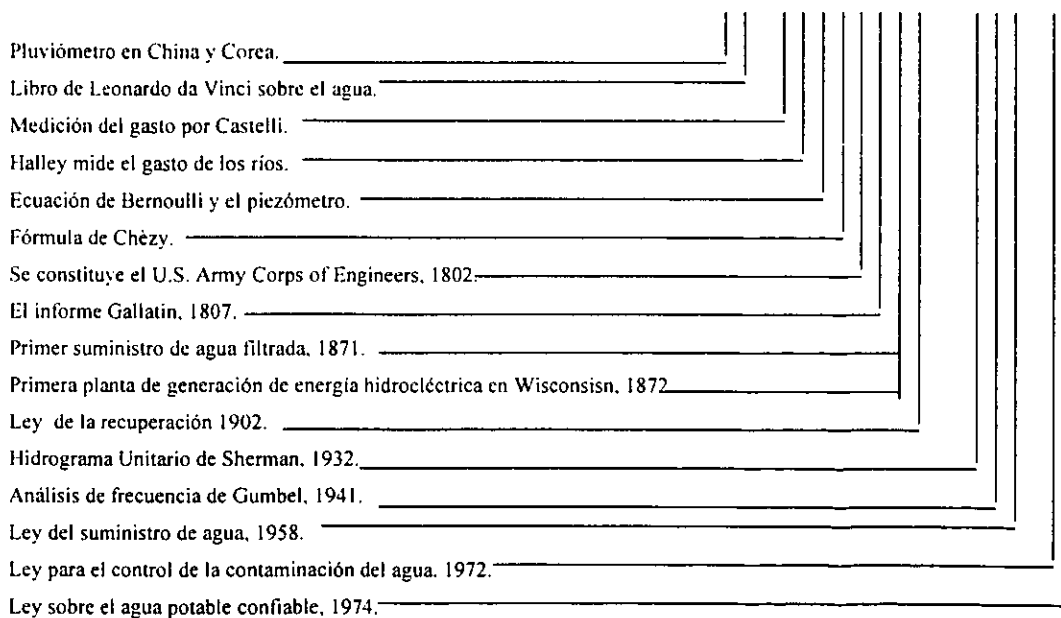
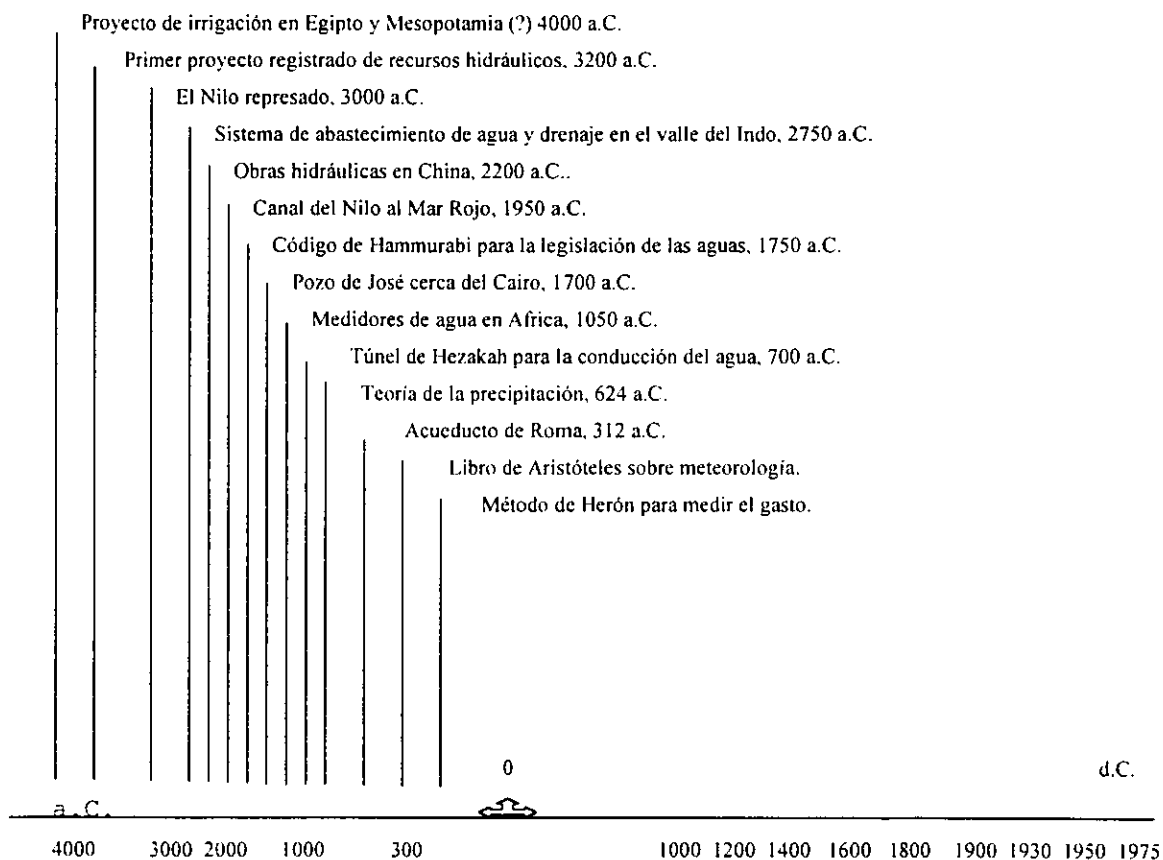
<sup>5</sup> Chézy, Antonio. Ingeniero francés nacido el 1/sept./1718 en Chanlons-sur- Marne, muerto en París el 4/oct./1798. Entró en el Cuerpo de Ingenieros de Puentes y Caminos en 1751. En 1771 era inspector general de París, en 1782 ayudante en la dirección de Puentes y Caminos y en la inspección de obras, en 1797 obtuvo el puesto en el Cuerpo de Puentes y Caminos: su gran mérito fue el de haber puesto las bases de la teoría de cantidad uniforme en los canales. Fuente: *Enciclopedia Italiana di Scienze, Lettere ed Arti*. Istituto della Enciclopedia italiana fondata da Giovanni Treccani (trad.M.Velasco V.) 38vols., Roma, Italia, 1950, pág.988.

Durante el s. XVIII hubo muchos avances, como por ejemplo, se desarrolló la fórmula para la medición de la fuerza de los fluidos, debida a Pitot, otro ejemplo se debe a Daniel Bernoulli quien desarrolló la ecuación de la energía, que es básicamente el principio de la conservación de la energía.

En 1807 el informe Gallatin, resultante de las expediciones de Lewis y Clark, representó el primer plan completo de recursos hidráulicos ideado en los Estados Unidos. En 1913 otro estudio debido a la inundación en el distrito comercial de Dayton, fué el de Morgan, que contemplaba mejoras en 63Km de diques y canales que atravesaban nueve ciudades y cinco embalses secos de detención de uso restringido; este se terminó en 1922.

En la siguiente página se muestra un línea de tiempo con los eventos más significativos de los recursos hidráulicos:

## LINEA DE TIEMPO DE LOS EVENTOS CULMINANTES DE LOS RECURSOS HIDRAULICOS



FUENTE: Helweg, Otto J. *Recursos Hidráulicos: Planeación y Administración*. Edit. Limusa. pág.16-17.

## 1.2 DEFINICIÓN BÁSICA

Hidrología

La hidrología estudia las aguas de la superficie terrestre, como las que fluyen en los ríos o las que están depositadas en los lagos y lagunas.

La importancia de lo anterior es la influencia del agua en la vida de los pueblos, por lo tanto, cuanto mayor sea el conocimiento científico que se tenga de las aguas superficiales mejor será el aprovechamiento de ellas.

Dentro de la hidrología uno de los principales objetivos es el de recopilar datos hidrológicos y meteorológicos, para generar la información necesaria para evaluar, desarrollar y administrar los recursos hídricos de una localidad así como el medio ambiente vinculado al recurso agua.

### 1.2.1 DEFINICIÓN Y OBJETO DE LA RED HIDROLÓGICA

Una red es un conjunto de estaciones hidrométricas y climatológicas ubicadas en un área específica.

"El objeto de una red es suministrar una densidad y distribución de estaciones en una región tal que, por interpolación entre conjuntos de datos provenientes de distintas estaciones, será posible determinar con suficiente exactitud para fines prácticos las características de los elementos básicos hidrológicos y meteorológicos en cualquier lugar de la región."<sup>6</sup>

Hay seis técnicas básicas utilizadas conjuntamente con las redes para la elaboración de revistas, boletines, avisos, etc., estos incluyen los datos relacionados con el agua, y se realizan mediante:

- a) análisis cartográfico
- b) análisis de correlación y regresión
- c) modelos probabilísticos
- d) modelos estocásticos

- e) modelos determinísticos y
- f) regionalización de parámetros de los modelos.

Cada método tiene aplicación particular y la elección depende de las limitaciones de los datos disponibles y el tipo de problema en consideración.

Los tipos de datos a considerarse son aquellos que requieren una observación prolongada y que son de importancia para el diseño de las redes hidrológicas. La falta de esta información puede generar el retraso en proyectos como abastecimiento de agua, etc.

Los datos hidrológicos y climatológicos son de muy diversas clases, se mencionarán sólo algunos:

- a) precipitación y capa de nieve
- b) nivel y caudal de los ríos y nivel de los lagos y embalses
- c) evaporación y evapotranspiración
- d) transporte y depósitos de sedimentos
- e) calidad química del agua superficial
- f) temperatura del agua
- g) capa de hielo en los ríos, lagos y embalses
- h) densidad del desarrollo de plantas acuáticas en el lecho del río
- i) humedad del suelo
- j) agua subterránea

La calidad de los datos está sujeta a algunos aspectos que deben considerarse como son:

- a) Instrumentación adecuada: Se deben seleccionar instrumentos capaces de medir o registrar datos precisos de acuerdo a las condiciones en las que son operados.
- b) Cuidado en las observaciones.

- c) Control de calidad en el procesamiento de datos: Son esenciales los procedimientos adecuados de control de calidad durante el procesamiento de datos para minimizar errores en los registros finales.
- d) Programación de la preparación de los datos: La preparación de los datos debe analizarse inmediatamente para ajustar las irregularidades en las primeras observaciones.

### 1.2.2 DENSIDAD Y CONFIGURACIÓN DE LAS REDES

La configuración de una red debe planificarse de modo que la ubicación de las estaciones representen todas las características físicas de la región. Es imposible definir una densidad uniforme para todos los países. Los detallados estudios realizados en distintas regiones han demostrado que entre los factores más importantes que determinan una densidad óptima están:

- a) Las condiciones geográficas e hidrológicas, particularmente las variaciones espaciales de los regímenes pluviométricos e hidrológicos, y
- b) La naturaleza de la hidrografía, por ejemplo, la existencia de numerosas corrientes pequeñas o un escaso número de ríos de importancia.

"La red mínima o básica es aquella que evita serias deficiencias en el desarrollo y la gestión de los recursos hídricos a una escala compatible con el nivel general de desarrollo del país."<sup>7</sup>

Con la red mínima o básica, comienza el proceso de expansión de toda una red completa que ayude al máximo aprovechamiento de los recursos hídricos de una región o país. Una nota importante es que la red mínima no es adecuada para la elaboración de grandes proyectos, sino que sólo sirve para iniciar la planificación del desarrollo de los recursos hídricos de una región.

---

<sup>7</sup> ibid., 3:3-9



Las estaciones que conforman una red pueden dividirse en tres categorías:

- a) estaciones principales o de base.
- b) estaciones secundarias.
- c) estaciones para fines especiales.

### 1.2.3 ESTACIONES HIDROMÉTRICAS

#### 1.2.3.1 Principales

Estas estaciones, llamadas estaciones base o permanentes, suministran la información básica para los estudios estadísticos y por lo tanto deberán estar en funcionamiento continua e indefinidamente. Se sugieren observaciones de un periodo de 30 a 40 años para obtener estimaciones confiables, mientras que se necesitan más de 70 años de observaciones en áreas con precipitación extremadamente variable.

#### 1.2.3.2 Secundarias

Estas estaciones, llamadas a veces estaciones "sputnik" o satélite, proveen los datos básicos para interpolar la variabilidad espacial de los elementos hidrológicos. Deben funcionar durante un número limitado de años o el tiempo suficiente para establecer una buena correlación entre ellas o con las estaciones de base o con las características del terreno.

#### 1.2.3.3 Para fines especiales

Se pueden establecer estaciones para fines especiales tales como la investigación o el incremento de datos suministrados por la red de estaciones de base y secundarias, en alguna región o con algún propósito específico como pueden ser los requerimientos operacionales. Las estaciones de referencia proporcionan observaciones sobre datos hidrológicos y climatológicos conexos, la cantidad y duración de los registros es lo más importante de estas estaciones, éstas pueden pertenecer a esta categoría. La duración de las estaciones especiales es función del propósito para el cual fueron hechas.

### 1.3 FRECUENCIA Y HORARIO DE LAS OBSERVACIONES

La frecuencia con la cual se observa un elemento depende de la velocidad en que cambia o varía el valor del elemento, las observaciones en las estaciones climatológicas e hidrométricas deben hacerse a las horas sinópticas, como se especifica en el reglamento técnico de la Organización Meteorológica Mundial (OMM) No.49; cada 3 y 6 horas (0000, 0300, 0600, 0900, 1200, 1800, y 2100). Y en los informes<sup>8</sup> donde se vacían los datos debe estar la fecha (año, mes, día) y las horas de observación de todas las estaciones para que estos informes sean válidos o confiables.

Si se adelanta una hora en el horario de verano, se tratará de que las observaciones se sigan haciendo a la misma hora que se hacían antes de la modificación del horario.

### 1.4 CUENCAS

#### 1.4.1 DEFINICIÓN DE CUENCA

En general se da el nombre de cuenca a cualquier depresión de la litósfera, esté cubierta o no por agua. *La cuenca de un río* es toda el área drenada del río y sus tributarios. Se llama *Cuenca de un lago* a la depresión ocupada por las aguas del lago.

La extensión de las cuencas es muy variable y depende de la topografía del terreno. Entre dos cuencas contiguas existe una cresta de separación que recibe los nombres de divisorias o parteaguas.

En cada región natural es deseable contar con una *cuenca representativa*, especialmente donde se espera un gran desarrollo económico o los problemas hidrológicos son particularmente difíciles, en estas cuencas se permite el

---

<sup>8</sup> Ver en el anexo 1 dos ejemplos de informe donde se vacía el gasto y datos pluviométricos.

estudio simultáneo de precipitación y escurrimiento, ayudando a equilibrar las deficiencias de periodos cortos de observación. Estas cuencas son elementos de una red completa.

#### 1.4.2 TIPOS DE CUENCAS

Las cuencas pueden ser de cuatro tipos distintos:

- Exorreicas, cuando sus aguas se vierten en el mar.
- Endorreicas, cuando sus ríos desembocan en una laguna interior.
- Arreicas, si no tienen actualmente escurrimientos de agua, aunque los hayan tenido en el pasado o puedan tenerlos en el futuro, de acuerdo con los climas cambiantes.
- Criptorreicas, con ríos subterráneos.

La cuenca es un concepto orográfico-hidrológico y tiene un carácter unitario, por lo que ofrece una base conveniente para los estudios de la geografía regional.

En el siguiente capítulo se muestran algunas técnicas estadísticas que se usan en el análisis y predicción hidrológica, basadas en los fenómenos naturales, que son transformados a números y recopilados por medio de la red de estaciones hidrométricas.

## CAPÍTULO II

## MÉTODOS ESTADÍSTICOS EN ANÁLISIS Y PREDICCIÓN HIDROLÓGICA

## 2.1 MODELADO DE SISTEMAS HIDROLÓGICOS

La investigación hidrológica incluye el estudio de casos aislados y también la inspección estadística del conjunto de datos. Los análisis estadísticos abarcan el ajuste de los datos a distribuciones de frecuencia y a modelos teóricos por regresión o por métodos de análisis de series de tiempo. La validez de las relaciones deducidas es contrastada con datos independientes. Un análisis puede ser desarrollado de diferentes formas:

## ANÁLISIS

## DESARROLLO

Determinístico:	Sigue las leyes que describen procesos físicos.
Paramétrico:	El análisis se lleva a cabo por comparación de datos hidrológicos registrados en diferentes emplazamientos y no en forma simultánea.
Probabilístico:	Se analiza la frecuencia de ocurrencia de diferentes magnitudes de las variables.
Estocástico:	En forma conjunta son analizados el orden secuencial y la frecuencia de ocurrencia de las diferentes magnitudes.

La hidrología de las aguas superficiales se refiere principalmente al modelado de los ríos (ecuaciones que presentan aproximaciones al flujo de los ríos) y al modelado de las cuencas (ecuaciones que presentan aproximaciones a la relación entre precipitación y escurrimiento).

Existen varias formas de plantear el problema o fenómeno:

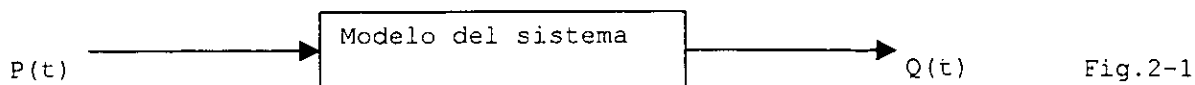
- Técnicas empíricas.
- Caja negra.
- Modelos genéticos o conceptuales.

Los modelos anteriores generan una única salida y sin probabilidad asociada de ocurrencia, por esto se encuentran referidos como **modelos determinísticos**.

"Los desarrollos en los modelados de sistemas hidrológicos se hallan estrechamente vinculados con el surgimiento de computadoras electrónicas y de técnicas para su aplicación. La disponibilidad de computadoras electrónicas y el desarrollo de métodos numéricos asociados han permitido a los hidrólogos llevar a cabo cálculos repetitivos complejos, usando para ello grandes cantidades de datos tal como lo necesitan el desarrollo y posterior aplicación de modelos de simulación determinísticos y estocásticos."<sup>9</sup>

*Modelos de caja negra (enfoque sistemático):* estos modelos no intentan modelar la estructura interna y la respuesta de la cuenca, sino solamente comparar la entrada y salida de la misma considerandolas equivalentes o iguales.

Por ejemplo, una cuenca de un río puede ser considerada como un sistema dinámico en el cual los factores de entrada son la precipitación y la fusión de las nieves y la salida es el caudal efluente. En forma de diagrama se puede representar como en la siguiente figura 2-1:



donde  $P(t)$  es la entrada y  $Q(t)$  es la salida, ambas funciones del tiempo  $h(t)$ .

---

<sup>9</sup> ibid., 5:1-3

*Modelos conceptuales:* estos modelos basan sus ecuaciones en varios conceptos que corresponden a procesos físicos de la formación de flujo. Uno de los aspectos relacionados con éstos es la calibración del modelo para ajustarlos a una cuenca; la mayoría de los parámetros se determinan por procesos iterativos, basados en datos históricos de entrada y salida.

Un ejemplo de modelo conceptual y de enfoque sistemático es el del Centro hidrometeorológico de la URSS; "... este modelo utiliza formulaciones conceptuales para deducir las pérdidas de la cuenca y luego aplica un enfoque sistemático para el routing de los volúmenes de escorrentía hasta la salida de la misma."<sup>10</sup>

Otros ejemplos son el modelo Sacramento y el modelo tanque, "... éste simula por medio de tanques puestos uno encima del otro, en donde la precipitación y la fusión de nieves se incorporan al tanque superior en el cual existe una salida inferior por donde escurre al siguiente tanque y así sucesivamente hasta el último tanque, el flujo que sale lateralmente va hacia el sistema de cauces, en el último tanque el flujo del fondo representa una pérdida para el sistema."<sup>11</sup>

En la siguiente página se muestra la figura 2-2 que representa gráficamente el modelo tanque.

También el término "Modelado de sistemas Hidrológicos" comprende en algunos casos el modelado estocástico, con el fin de simular series hidrológicas de tiempo. Haciendo notar la reproducción de las características estadísticas de tales series y no el modelado de las relaciones entrada-salida.

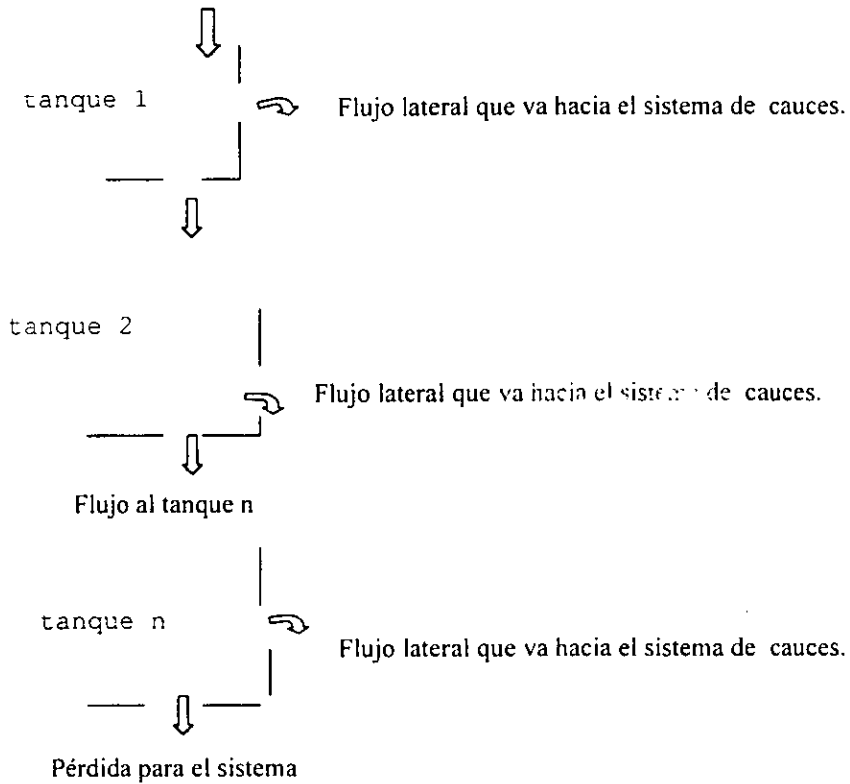
---

<sup>10</sup> ibid., 5:128

<sup>11</sup> ibid., 5:132-133

Figura 2-2 Modelo Tanque.

La precipitación y fusión de nieves se incorporan al primer tanque y así sucesivamente.



## 2.2 SIMULACIÓN ESTOCÁSTICA DE SERIES HIDROLÓGICAS DE TIEMPO

"Los modelos estocásticos son modelos de caja negra en los que los parámetros son estimados utilizando propiedades estadísticas de las series observadas....El modelado estocástico ha sido del mismo modo utilizado en el establecimiento de límites de confianza (probabilidades) en predicciones de flujo en tiempo real y en la realización de datos de precipitación para uso en modelos determinísticos".<sup>12</sup>

<sup>12</sup> *ibid.*, 5:135

### 2.2.1 MODELOS MARKOVIANOS LAG-1

Muchos modelos utilizados en la simulación de volúmenes de flujos mensuales, estacionales o anuales tienen estructura Markoviana de primer orden; es decir que el flujo durante un periodo dado está determinado por el valor del mismo en el periodo anterior y un impulso aleatorio. Un ejemplo de modelos para flujos mensuales puede ser expresado según:

$$Q_i = \bar{Q}_j + \rho_j (\sigma_j / \sigma_{j-1}) (Q_{i-1} - \bar{Q}_{j-1}) + \varepsilon_i \sigma_j \sqrt{1-\rho_j^2} \quad (\text{ecuación 2.1})$$

Donde:

$Q_i$  es el flujo del  $i$ -ésimo miembro de las series numeradas consecutivamente desde 1 haciendo caso omiso del mes o año;

$j$  es el mes en el cual se ubica el  $i$ -ésimo miembro de la serie;

$\bar{Q}_j$  es el flujo medio correspondiente al  $j$ -ésimo mes;

$\sigma_j$  es la desviación típica correspondiente al  $j$ -ésimo mes;

$\rho_j$  es el coeficiente de correlación serial entre  $Q_j$  y  $Q_{j-1}$  ;

$\varepsilon_i$  es una variable aleatoria generada mediante una distribución apropiada con un valor medio nulo, varianza unitaria e independencia serial.

La ecuación 2.1 es adecuada para flujos estacionales ( $j=1,2,3,\text{ó }4$ ) y flujos anuales ( $j=1$ ). En el último de los casos la ecuación 2.1 se transforma en:

$$Q_i = \bar{Q} + \rho(Q_{i-1} - \bar{Q}) + \varepsilon_i \sigma \sqrt{1-\rho^2} \quad (\text{ecuación 2.2})$$

se supone que los valores de  $\bar{Q}$ ,  $\rho$  y  $\sigma$ , derivados de registros históricos aplicables para este fin, necesitándose solamente seleccionar un valor inicial de  $Q_{i-1}$  para simular una serie de longitud requerida. Generalmente se utilizan las técnicas Montecarlo con valores secuenciales de la variable aleatoria derivados mediante computadora.

El desarrollo y aplicación de los modelos representados en la ecuación 2.1 son relativamente directos y simples.



### 2.2.2 MODELOS AUTORREGRESIVOS DE PROMEDIOS MÓVILES INTEGRADOS (ARIMA)

La cantidad más importante de modelos estocásticos univariados está representada en el grupo ARIMA desarrollado por Box y Jenkins. Existen tres clases: autorregresivo (AR), promedios móviles (MA) y compuesto (ARMA). Las primeras dos clases (de orden  $p$  y  $q$ ) son respectivamente:

$$x_t = \varepsilon_t + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \dots + \Phi_p x_{t-p} \quad (\text{ecuación 2.3})$$

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (\text{ecuación 2.4})$$

donde  $x_t$  es la desviación de la observación  $t$ -ésima respecto al promedio de las series,  $\Phi_i$  y  $\theta_i$  son parámetros que deben ser estimados y  $\varepsilon_t$  es una variable aleatoria definida de acuerdo con lo antes mencionado. La tercer clase (ARMA) es una combinación de las dos primeras conteniendo todos los términos no repetitivos de las ecuaciones 2.3 y 2.4:

El ajuste de los modelos ARIMA puede llevarse a cabo por medio del siguiente procedimiento sistemático:

- a) identificación - Se compara la correlación de las series a estudio con las funciones de autocorrelación de diferentes modelos ARIMA a los efectos de la apropiada selección de la clase y del orden;
- b) estimación - Los parámetros del modelo se estiman mediante el uso iterativo de mínimos cuadrados, con la restricción de que los residuos se distribuyan independientemente y la suma de los cuadrados sea mínima;
- c) verificación - El carácter de aleatorio de los residuos se verifica para comprobar la adecuación del modelo seleccionado.

Los modelos ARIMA se utilizan para generar secuencias sintéticas de flujo mediante técnicas Montecarlo en la forma previamente descrita. Sin embargo, los métodos estocásticos de generación deben emplearse con precaución.

### 2.2.3 MODELOS FRACCIONALES GAUSSIANOS DE RUIDO Y DE PROCESO DE LÍNEA QUEBRADA

Hurst, H.E. descubrió que los registros geofísicos muy extensos exhiben características que no concuerdan con procesos estacionarios autorregresivos Markovianos. Este descubrimiento condujo al desarrollo de dos modelos estocásticos que pueden acomodar elementos de persistencia a largo plazo (frecuencia baja).

El primero de ellos se denomina usualmente modelo *fraccional gaussiano de ruido (FGN)*. La base de este modelo es un proceso aleatorio autosimilar, caracterizado por una función de densidad espectral que subraya frecuencias muy bajas tipificadas por el fenómeno Hurst. Así mismo se demuestra que un modelo de memoria extensa de un proceso de línea quebrada preservará el fenómeno Hurst. Los hallazgos de Hurst no indican necesariamente una persistencia a muy largo plazo y, mas aún, algunas versiones de los modelos ARIMA son capaces de simular sustanciales efectos en baja frecuencia. La no estacionalidad del valor medio del proceso podría derivar en las características que Hurst halló al analizar registros largos, ya sea que fuera el resultado de cambios climáticos, actividades humanas o simple falta de homogeneidad de la serie de datos.

### 2.3 PREDICCIÓN HIDROLÓGICA

Según la definición del diccionario *predecir* es conocer el futuro por conjeturas o indicios, pero ¿Qué es una predicción hidrológica?

"Una predicción hidrológica es la estimación previa del estado futuro de los fenómenos hidrológicos." <sup>13</sup>

Estos fenómenos se refieren a las lluvias, que sirven para el riego, el abastecimiento en las presas, las crecidas de los ríos y todo lo relacionado con la distribución del agua en la tierra.

---

<sup>13</sup> *ibid.*, 6:1

La necesidad de predecir es necesaria para el aprovechamiento de los recursos de una región, como por ejemplo el abastecimiento de agua, para evitar incendios en los bosques, etc. Generalmente los errores en las predicciones se deben a errores en la medición o al modelo utilizado, o a los datos que varían rápidamente.

### 2.3.1 VARIABLES USUALES EN UNA PREDICCIÓN

Los elementos básicos de un régimen hídrico, que son de interés práctico para la predicción:

- a) Volumen de escorrentía en diferentes periodos de tiempo (por ejemplo, periodos de aguas altas o bajas, mes, estación, año).
- b) Hidrograma de caudales o niveles.
- c) Nivel o caudal de punta crecida y su tiempo de ocurrencia.
- d) Niveles máximos del agua en los lagos y fechas en que se alcanzarán dichos niveles.
- e) Niveles medios y mínimos de ríos navegables y de lagos durante varios periodos del calendario.
- f) Altura de olas originadas por viento en lagos y grandes embalses.
- g) Elevación del nivel del agua por causa del viento en lagos y aguas costeras.
- h) Parámetros de salida de agua tales como temperatura y turbidez.

En el régimen de hielo en ríos, lagos y embalses:

- a) Fecha de otoño en que comienza la formación de hielo flotante.
- b) Fecha en que se cubren totalmente de hielo.
- c) Espesor de la capa de hielo.
- d) Fecha de primavera en que comienza a fragmentarse el hielo.
- e) Fecha en que el hielo desaparece completamente de lagos y embalses.

Y en el régimen de aguas subterráneas y del agua del suelo:

- a) Máximo agotamiento y mes en que se produce.
- b) Fecha de recuperación de la capacidad de campo.

- c) Fecha de cese de flujo de manantiales.
- d) Nivel máximo o futuro en el acuífero.

### 2.3.2 COMPONENTES DE UNA PREDICCIÓN

Las predicciones hidrológicas tienen seis características principales:

- a) La variable que se debe predecir, elemento del régimen que se debe predecir.
- b) Plazo, también llamado periodo de predicción o de advertencia.
- c) El método de computación usado.
- d) La finalidad de la predicción.
- e) La forma de presentación, por ejemplo, valor único esperado, hidrograma total, distribución probabilística, etcétera.
- f) Los métodos de difusión.

### 2.3.3 TÉCNICAS USADAS EN PREDICCIÓN

Las técnicas, clasificadas en las cuatro categorías que se mencionan a continuación, deben ser juzgadas en relación con su aplicación en situaciones de predicción en tiempo real y no de su potencial en la forma pasada del análisis, apropiada para el diseño de modelos. Las cuatro categorías son:

- a) Modelos empíricos y modelos conceptuales de base física de los procesos hidrológicos en la fase terrestre.
- b) Métodos basados en procesos de intercambio de calor dentro de los cuerpos de agua.
- c) Modelos empíricos y modelos físicos para calcular el movimiento y la traslación de la onda de crecida.
- d) Métodos basados en el análisis de los procesos de circulación atmosférica.

Las técnicas de correlación y regresión, análisis de series de tiempo, técnicas de ajuste de predicción y predicción probabilística son técnicas matemáticas de propósitos generales que pueden aplicarse en cualquiera de los métodos antes mencionados.

### 2.3.3.1 Correlación y regresión

El coeficiente de correlación mide la unión entre dos variables y es una herramienta matemática ampliamente usada en numerosos análisis hidrológicos. La regresión es una extensión del concepto de correlación que proporciona formulas para deducir una variable de interés, por ejemplo, caudal de estiaje estacional a partir de una o más observaciones disponibles, por ejemplo, máximo nivel invernal de aguas subterráneas.

La fórmula para calcular el coeficiente de correlación  $r$  entre  $n$  valores de  $x$  &  $y$  es:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (\text{ecuación 2.5})$$

donde:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad \bar{y} = \sum_{i=1}^n \frac{y_i}{n}$$

La falta de correlación no implica falta de unión. Como  $r$  indica solo una relación lineal, una relación curvilínea estricta no se verá necesariamente reflejada por un alto valor de  $r$ . Recíprocamente, una correlación entre dos variables no garantiza que tengan unión causal. Una simple relación de datos dispersos entre dos variables de interés conduce a una correlación gráfica.

Cuando tanto  $x$  como  $y$  tienen una estructura de series de tiempo, especialmente una tendencia, se hará lo necesario para separar esa estructura

antes de correlacionar, y se tendrá mucha cautela en la interpretación de la prueba de significación.

Las ecuaciones de regresión tienen muchas aplicaciones en hidrología. Su forma general es:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + e_t \quad (\text{ecuación 2.6})$$

donde las X son las variables observadas y Y es el futuro valor de la variable que se debe predecir. Los b son los coeficientes de regresión estimados a partir de los valores observados de X e Y.

Las variables X pueden incluir nivel o caudal aguas arriba, lluvia, condiciones de la cuenca, temperatura o lluvia estacional.

La variable Y puede ser el nivel máximo o mínimo. El coeficiente de correlación múltiple mide el grado de ajuste en la relación. Otra medida del ajuste, el error estandar o estimado, mide la desviación típica de las diferencias del conjunto de calibración respecto a la línea de regresión.

### 2.3.3.2 Análisis de series de tiempo

Como la escorrentía es un indicador de la situación de la cuenca, el análisis de series de tiempo puede ser usado para establecer relaciones de predicción. Un ejemplo de aproximación consiste en el uso de los modelos ARMA (promedio autorregresivo móvil) que se ajustan para cuencas con pocos datos de precipitación, por que solo es necesario el caudal anterior para hacer predicciones del tipo:

$$Q_{t+1} = a_0Q_t + a_1Q_{t-1} + a_2Q_{t-2} + \dots + b \quad (\text{ecuación 2.7})$$

donde  $Q_{t+1}$  es la predicción con un tiempo de preaviso unitario y  $Q_{t-1}$  son los valores medidos y periodos anteriores. Los coeficientes  $a_i$  y b son estimados en el análisis de series de tiempo. Además del valor  $Q_{t+1}$  predicho, un modelo de series de tiempo brinda la distribución de posibles desviaciones del mismo respecto a la predicción, por lo que se puede disponer de información sobre el error de predicción.

Para que una predicción de caudales mensuales por series tiempo sea confiable, la autocorrelación en las series de tiempo mensuales debe ser

grande. Es este el caso de ríos extensos y de cursos de agua que drenan a grandes acuíferos de agua subterránea. Sin embargo, las predicciones sólo serán factibles con plazos de 1 a 4 meses de adelanto.

Se pueden incluir variables climáticas de entrada en un modelo de series de tiempo, si estos datos están disponibles, a menudo será preferible hacer predicciones usando métodos de regresión o utilizando modelos conceptuales.

### 2.3.3.3 Técnicas de ajuste de la predicción.

Si se hace una observación de un valor predicho  $Y_1$ , entonces existe la oportunidad de ajustar las predicciones con un error conocido  $e_1 = Y_1 - \hat{Y}_1$ , donde  $\hat{Y}_1$  es la estimación prevista. De la forma más simple, el ajuste consiste en la suma del error  $e_1$  a la nueva predicción  $\hat{Y}_{1+1}, \hat{Y}_{1+2}, \dots$ . Para evitar discontinuidades, el ajuste es normalmente incluido en el hidrograma calculado sobre varios periodos de tiempo.

Un procedimiento más complicado consiste en someter la serie de errores,  $e_1, e_2, \dots, e_i$  a un análisis de series de tiempo, para extraer posibles tendencias o periodicidades que pueden ser extrapoladas para estimar el nuevo error  $\hat{e}_{1+1}$  que se sumará a la nueva predicción  $\hat{Y}_{1+1}$ .

Los valores  $Y_1, Y_2, \dots, Y_i$ , se pueden usar para redefinir las variables de estado del modelo de predicción. Esto se denomina estimación repetitiva, y si el modelo obtenido es muy sencillo pero representativo, proporciona la base de una estrategia óptima y formas para ajustar los valores producidos por el modelo.

"El filtro Kalman y el filtro Kalman extendido son técnicas de estimación repetitiva que se han aplicado a la predicción. Estas técnicas requieren una habilidad matemática e hidrológica considerable para asegurar que el modelo de predicción este en una forma conveniente para el análisis."<sup>14</sup>

---

<sup>14</sup> ibid., 6: 35

#### 2.3.3.4 Predicción probabilística

Las predicciones a largo plazo, (especialmente de escorrentía estacional), generalmente están expresadas en términos probabilísticos o sea una distribución estadística de posibles volúmenes de escorrentía dependientes de la lluvia posterior a la fecha en que se hizo la predicción. Una fuente de incertidumbre es el estado del tiempo futuro entre la fecha de preparación de la predicción y la fecha de ejecución de la misma. Como ejemplo, si una predicción basada en una regresión da:

$$Q_{\text{verano}} = b_0 + b_1 R_{\text{otoño}} + b_2 R_{\text{invierno}} + b_3 R_{\text{primavera}} + b_4 R_{\text{verano}} \quad (\text{ecuación 2.8})$$

Puede darse una predicción probabilística luego de recibir los datos de lluvia del otoño e invierno anteriores, teniendo en cuenta la distribución de posibles lluvias que puedan ocurrir en primavera y verano.

A menos que el modelo de predicción sea muy sencillo, será necesario simular posibles valores  $Q_{\text{verano}}$  ya sea mediante repetidos muestreos a través de la distribución de valores  $R_{\text{primavera}}$  y  $R_{\text{verano}}$ , o sino mediante la aplicación del modelo o la serie histórica de valores observados durante el periodo registrado.

Si se usa el enfoque histórico, es deseable tener al menos 30 años de registros para obtener un conjunto representativo de combinaciones. La aplicación de esta técnica no está limitada, puede aplicarse a datos reales o sintéticos para obtener una distribución de resultados posibles.

#### 2.3.3.5 Desarrollo del procedimiento de predicción

Antes de desarrollar un procedimiento de predicción, debe establecerse una clara definición del objetivo de la misma. Se identificarán las variables a predecir, el tipo de predicción, plazo de predicción, la predicción o aviso hidrológico puede ser presentada en las siguientes formas:

- a) Un solo valor de la variable predicha (ejemplo, altura de cresta, caudal de punta, fecha de congelación).



- b) Una variable predicha como función del tiempo (ejemplo, el hidrograma de caudales o de alturas).
- c) Una distribución probabilística de la variable predicha expresando el campo de posibles valores futuros en términos de su probabilidad de ser excedidos, teniendo en cuenta las condiciones futuras.
- d) Una definición del intervalo de valores de una variable predicha, en relación con el valor medio o normal, por ejemplo, <<por encima del promedio>> o <<muy por debajo de lo normal para la estación>>.
- e) Un mensaje que intenta advertir a todos los afectados sobre la próxima ocurrencia de acontecimientos hidrológicos peligrosos como crecidas repentinas, onda de borrasca, congelación de un lago o de un tramo de río.

Generalmente, los procedimientos se hacen más complejos cuando se dispone de tecnología avanzada. Cuando se usan simples instrumentos manuales y computadoras pequeñas es apropiado que se usen métodos basados en correlaciones o índices pues es limitada la capacidad de operación de archivos históricos.

#### 2.3.3.6 Determinación de parámetros y coeficientes del modelo

La mayoría de los procedimientos de predicción incluyen parámetros numéricos, cuyos valores son determinados a partir de datos durante la etapa de desarrollo de la predicción.

Para muchos procedimientos de predicción y especialmente aquellos que usan modelos de base física, el proceso de prueba contiene elementos subjetivos. Los pasos para probar modelos de base física son:

- a) Decisión en cuanto a la subdivisión de la cuenca.
  - b) Estimación inicial de los parámetros.
  - c) Corridas de simulación por aproximaciones sucesivas para mejorar la estimación de parámetros.
  - d) Optimización automática.
- a) En la etapa inicial de estimación los parámetros que controlan el almacenamiento en el suelo como la infiltración, la recesión, pueden ser

inferidos a partir del tipo de suelo de la cuenca, complementándose con las características generales de respuesta de la cuenca.

b) Las estimaciones iniciales se usan para simular el hidrograma continuo. Para perfeccionar la predicción se ajustan manualmente los valores de los parámetros por aproximaciones sucesivas. El proceso de prueba debe concentrarse en el elemento que se va a predecir.

c) El grado de ajuste entre los valores observados y simulados  $x$  del elemento de predicción se mide mediante la función objetivo. La forma de función más común es  $F = \Sigma(x - \bar{y})^2$ , aunque otras pueden ser más apropiadas para fines particulares. En una predicción con fines generales, los valores  $x$  &  $y$  pueden referirse a las ordenadas del hidrograma. El fin de la prueba es encontrar los valores óptimos de los parámetros sujetos a cualquier restricción que pueda existir en el intervalo de validez del modelo y de valores de los parámetros. Una alternativa de  $F$  es la eficiencia;  $R^2 = (V-F)/V$  donde  $V = \Sigma(y - \bar{y})^2$  es la suma de los cuadrados de las diferencias respecto a la media de las observaciones. Entonces el proceso de prueba es la búsqueda del mayor valor de  $R^2$ .

d) Luego de disponer de valores iniciales aceptables para todos los parámetros se puede usar la optimización automática. Solo pocos parámetros pueden ser ajustados mediante esta optimización. La cantidad de ajuste de estos parámetros puede ser limitada por restricciones basadas en las corridas de aproximación sucesiva. La optimización automática incluye procedimientos de cálculo numérico.

#### 2.3.3.7 Criterios de verificación

Un método de predicción puede considerarse como eficaz y aplicarse, si los errores de predicción son sustancialmente menores a los que pueden resultar de una predicción ingenua. El *valor medio* es tal vez la más simple de las predicciones ingenuas, otra opción, como la *hipótesis* de que la variable predicha permanecerá siendo la misma que la observada en el momento inicial de la predicción (llamada predicción de inercia), o que la tendencia continuará en un futuro inmediato.

La evaluación estadística de la efectividad de cualquier método de predicción debe basarse en una comparación con resultados obtenidos aplicando una de estas predicciones ingenuas. El coeficiente de determinación  $I_d$  mide la efectividad de la predicción comparada con el uso de la media

$$I_d = 1 - \frac{S_e^2}{S_y^2} \quad (\text{ecuación 2.9})$$

donde  $S_e$  es el error típico del método, y  $s_y$  es la desviación típica de la variable predicha. Una comparación con una predicción de inercia puede lograrse haciendo que  $s_y$  en la ecuación anterior sea el error típico de la predicción de inercia. Se puede usar un parámetro estadístico similar al coeficiente de determinación para evaluar la efectividad de ciertos métodos de predicción a corto plazo.

$$I_\Delta = 1 - \frac{S_e^2}{S_\Delta^2} \quad (\text{ecuación 2.10})$$

donde  $S_\Delta^2$  es la desviación cuadrática media de las desviaciones de la variable predicha respecto a su variación media durante el periodo cubierto por la predicción.

El valor está dado mediante la fórmula:

$$S_\Delta^2 = \frac{\sum_1^N (\Delta y_j - \overline{\Delta y})^2}{N} \quad (\text{ecuación 2.11})$$

Aquí  $\Delta y = y_t - y_{t+1}$  es la variación de la variable predicha durante el plazo de predicción  $t$  (esto es un error de predicción de inercia) y  $\overline{\Delta y}$  es la variación media. Los valores  $S_e$ ,  $S_y$  y  $S_\Delta$  deben ser calculados teniendo en consideración sus variaciones estacionales. Los coeficientes  $I_d$  &  $I_\Delta$  son características estadísticas completamente confiables de las variables  $y$  y de los errores en el caso de una distribución normal.

Cuando la distribución no es normal  $I_d$  &  $I_\Delta$  sólo pueden ser usados después de confirmada la semejanza de las distribuciones empíricas de error típico de ambas variables (la predicción y el método de referencia). La expresión del valor del error típico es:

$$z = (x - \bar{x})/s_x \quad (\text{ecuación 2.12})$$

donde  $z$  es la variable normalizada;  $x$  la variable aleatoria en cuestión;  $\bar{x}$  es el valor medio de  $x$  y  $s_x$  la desviación típica de  $x$ . Luego la normalización del valor medio de la variable se hace cero y la desviación típica uno. La semejanza entre la distribución empírica del error típico y la distribución normal puede verificarse mediante los métodos de Kolmogorov-Smirnov y Chi cuadrado.

Si se encuentra una diferencia significativa entre ambas distribuciones, se concluye que los coeficientes  $I_d$  &  $I_\Delta$  proporcionan solamente una estimación imperfecta de la verdadera efectividad de la predicción. Los coeficientes  $I_d$  &  $I_\Delta$  definen la confiabilidad del método de predicción:

- Si  $I_d(I_\Delta) = 1$  el método es perfecto y libre de errores.
- Si  $I_d(I_\Delta) = 0$  el método no proporciona información adicional por encima de la predicción ingenua.
- Si  $I_d(I_\Delta) < 0$  el método genera mayores errores que la predicción ingenua.

Teniendo en cuenta los errores probables en  $I_d(I_\Delta)$ , estos índices deberán satisfacer la siguiente desigualdad para que el método tenga algún valor:

$$\max(I_d, I_\Delta) \geq I_{cr}^N \quad (\text{ecuación 2.13})$$

donde  $I_{cr}^N$  es el valor crítico de los índices dependiente de  $N$ , siendo  $N$  el número total de errores de predicción. Los valores críticos de  $I_{cr}^N$  se dan a continuación:

N	5	10	20	50	100
$I_{cr}^N$	0,0870	0,648	0,497	0,375	0,316

El criterio de la desigualdad significa que con un nivel de confianza del 95% el método de predicción verificado asegura al menos una reducción del 10 por ciento en el error cuadrático medio de la predicción con respecto a la desviación típica de una variable predicha.

En el siguiente capítulo se presenta la base teórica de las técnicas elegidas (regresión lineal y series de tiempo) para su posterior aplicación.

## CAPÍTULO III

### TÉCNICAS ESTADÍSTICAS

#### 3.1 REGRESIÓN LINEAL

Dada una variable dependiente y un conjunto de una o más variables independientes, todas ellas cuantitativas, la regresión lineal consiste en obtener una función lineal de las variables independientes que permitan explicar o predecir el valor de la dependiente.

A partir de  $(y_1, x_{11}, \dots, x_{1p}), \dots, (y_n, x_{n1}, \dots, x_{np})$ , muestra de  $n$  observaciones de las variables  $Y, X_1, \dots, X_p$ , se trata de aproximar los valores de  $Y$ , variable dependiente, mediante una función de las variables  $X_1, \dots, X_p$ , variables independientes, que exprese la asociación lineal entre  $Y$  y  $X_1, \dots, X_p$ :

$$Y = \beta_1 X_1 + \dots + \beta_p X_p + \beta_0 + e \quad (\text{ecuación 3.1})$$

donde  $\beta_0, \dots, \beta_p$  son parámetros desconocidos a estimar y  $e$  es una variable error  $N(0, \sigma^2)$ . En particular, para cada observación se tendrá:

$$y_i = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \beta_0 + e_i \quad i=1, \dots, n \quad (\text{ecuación 3.2})$$

donde los  $e_i, i = 1, \dots, n$ , son independientes entre sí.

##### 3.1.1 COEFICIENTE DE CORRELACIÓN SIMPLE

El coeficiente de correlación lineal simple,  $\rho$ , mide el grado de asociación lineal entre dos variables medidas en escala de intervalo o de razón, tomando valores entre -1 y 1. Valores de  $\rho$  próximos a 1 indicarán fuerte asociación lineal positiva (a medida que aumentan los valores de una de las dos variables, aumentan los de la otra); valores de  $\rho$  próximos a -1 indicarán fuerte asociación lineal negativa (a medida que aumentan los valores de una de las dos variables, disminuyen los de la otra), y valores de  $\rho$  próximos a 0 indicarán no asociación lineal (lo que no significa que

no pueda existir otro tipo de asociación). El estimador muestral para  $\rho$  es el coeficiente de correlación muestral,  $r$ .

El coeficiente de correlación es una medida del grado de asociación lineal que depende del tamaño muestral: un mismo valor del coeficiente de correlación muestral, calculado a partir de muestras de distinto tamaño de dos pares de variables, no corresponde a un mismo grado de asociación lineal. Para determinar si la asociación es estadísticamente significativa, se puede plantear la hipótesis nula de que el coeficiente de correlación lineal es igual a cero:

$$H_0: \rho = 0$$

El estadístico de contraste se construye a partir del coeficiente de correlación muestral,  $r$ , y del tamaño de la muestra,  $n$ . Si el p-valor asociado es menor que  $\alpha$ , se rechazará la hipótesis nula al nivel de significación  $\alpha$ .

### 3.1.2 REGRESIÓN LINEAL SIMPLE

En el caso de una única variable independiente,  $X$ , se habla de regresión lineal simple. La correspondiente ecuación de regresión será del tipo:

$$Y = \beta_1 X + \beta_0 + e \quad (\text{ecuación 3.3})$$

y, en particular, para cada observación:

$$y_i = \beta_1 x_i + \beta_0 + e_i \quad i = 1, \dots, n \quad (\text{ecuación 3.4})$$

donde  $\beta_0$  y  $\beta_1$  son parámetros desconocidos a estimar y los  $e_i$  proceden de variables independientes, normales, de media 0 y la misma varianza.

### 3.1.2.1 Estimación de los parámetros

El criterio para obtener los coeficientes de regresión  $B_0$  y  $B_1$ , estimaciones de los parámetros desconocidos  $\beta_0$  y  $\beta_1$  respectivamente, es el de mínimos cuadrados, que consiste en minimizar la suma de los cuadrados de los residuos. Si  $\hat{y}_i$  es la estimación de  $y_i$  mediante el modelo de regresión lineal:

$$\hat{y}_i = B_1 x_i + B_0 \quad (\text{ecuación 3.5})$$

el residuo correspondiente,  $E_i$  será la desviación de cada observación al valor estimado:

$$E_i = y_i - \hat{y}_i \quad (\text{ecuación 3.6})$$

### 3.1.2.2 Análisis de los residuos

En la regresión lineal se supone que los verdaderos errores,  $e_i$ , son independientes con distribución  $N(0, \sigma^2)$ . Los residuos,  $E_i$ , son las estimaciones de los verdaderos errores, y la estimación de  $\sigma^2$  es la media de los cuadrados de los residuos,  $s^2$ , donde  $s$  es el error típico de la estimación. Si el modelo ajustado es apropiado, los residuos deben presentar características similares.

El hecho de que la media de los residuos sea igual a cero es consecuencia del método de estimación de los parámetros de la función de regresión. Respecto a la normalidad, la distribución de la variable formada por los residuos debe ser Normal: los residuos observados y los esperados bajo hipótesis de distribución Normal deben ser parecidos. para comprobarlo, una alternativa es el gráfico de probabilidad normal, que permite comparar, gráficamente, la función de distribución observada en la muestra con la función de distribución de una Normal(0,1) (por lo que la variable objeto de análisis debe tener media 0 y desviación típica 1).



### 3.1.2.2.1 Prueba Durbin-Watson

La prueba Durbin-Watson para autocorrelación consiste en determinar si o no, el parámetro de autocorrelación  $\rho$  es cero en  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$  (ec. 3.7)

$$Y \quad \varepsilon_t = \rho \varepsilon_{t-1} + u_t \quad (\text{ec. 3.8})$$

Donde  $\rho$  es un parámetro tal que  $|\rho| < 1$  y  $u_t$  es independiente  $N(0, \sigma^2)$ .

Notar que si  $\rho=0$ , entonces  $\varepsilon_t = u_t$  de aquí el error  $\varepsilon_t$  es independiente cuando  $\rho=0$  dado que el término de desorden  $u_t$  es independiente.

La prueba considera dos alternativas  $H_0: \rho=0$  y  $H_a: \rho>0$ . El estadístico D de Durbin-Watson se obtiene usando mínimos cuadrados para especificar la función de regresión, calculando los residuales :

$$e_t = y_t - \hat{y}_t \quad (\text{ec. 3.9}) \quad \text{y entonces calculando el estadístico D:}$$

$$D = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (\text{ecuación 3.10})$$

donde  $n$  es el número de casos.

Durbin y Watson establecieron límites superior  $d_u$  e inferior  $d_l$  para valores críticos; tal que un valor D fuera de los límites nos lleva a tomar una decisión, la regla es la siguiente:

Si  $D > d_u$  aceptamos  $H_0$

Si  $D < d_l$  aceptamos  $H_a$

Si  $d_u \leq D \leq d_l$  la prueba es poco convincente

Pequeños valores de D nos llevan a concluir que  $\rho > 0$  por que el error contiguo  $\varepsilon_t$  y  $\varepsilon_{t-1}$  tienden a ser de la misma magnitud cuando están positivamente autocorrelados, de aquí las diferencias en los residuales  $e_t$  y  $e_{t-1}$  podrían tender a ser pequeñas cuando  $\rho > 0$  llevándonos a un pequeño numerador en D y por lo tanto a un pequeño estadístico D. Éste estadístico D mide el grado de autocorrelación entre el residuo correspondiente a cada observación y la anterior. Si su valor es próximo a 2, los residuos estarán

incorrelados, si se aproximan a 4, estarán negativamente autocorrelados, y si se aproxima a 0, estarán positivamente autocorrelados.

### 3.1.2.3 Análisis de la varianza

El análisis de la varianza se basa en que la variabilidad total de la muestra puede descomponerse en la variabilidad explicada por la regresión y la variabilidad residual:

$$SC_{total} = SC_{reg} + SC_{res} \quad (\text{ecuación 3.11})$$

donde:

$SC_{total}$ : mide las desviaciones de las observaciones,  $y_i$ , a la media muestral de  $Y$ .

$SC_{reg}$  mide las desviaciones de las estimaciones mediante el modelo de regresión lineal,  $\hat{y}_i$ , a la media muestral de  $Y$ .

$SC_{res}$  mide las desviaciones de las observaciones,  $y_i$ , a las estimaciones mediante el modelo de regresión lineal,  $\hat{y}_i$ .

La tabla de análisis de la varianza se construye a partir de esta descomposición y proporciona el estadístico  $F$  que permite contrastar la hipótesis nula de que la pendiente de la recta de regresión es igual a cero:

$$H_0: \beta_1 = 0$$

Tabla: Análisis de la varianza para el modelo de regresión lineal

Fuente de variación	Suma de cuadrados	Grados de libertad	Medias de cuadrados	Estadístico F
Regresión	$SC_{reg}$	$p$	$MC_{reg} = \frac{SC_{reg}}{p}$	$F = \frac{MC_{reg}}{MC_{res}}$
Residual	$SC_{res}$	$n-p-1$	$MC_{res} = \frac{SC_{res}}{n-p-1}$	
Total	$SC_{total}$	$n-1$		

Además, se verifica que:

$$r^2 = \frac{SC_{reg}}{SC_{tot}} \quad (\text{ecuación 3.12})$$

donde  $r^2$  es el cuadrado del coeficiente de correlación muestral, estimación del cuadrado del coeficiente de correlación,  $\rho^2$ , al que se denomina coeficiente de determinación.

En consecuencia, el coeficiente de determinación puede interpretarse como la proporción de variabilidad total de la variable dependiente explicada mediante la recta de regresión. La hipótesis será entonces equivalente a la hipótesis:

$$H_0: \rho^2 = 0$$

o lo que es lo mismo, a la hipótesis de que Y & X están incorreladas.

Observar que, según la expresión del estadístico F, cuanto mayor sea su valor mejor será la predicción mediante el modelo lineal respecto a la predicción mediante la media muestral. Si el p-valor asociado a F es menor que  $\alpha$ , se rechazarán las dos hipótesis nulas planteadas al nivel de significación  $\alpha$ .

### 3.1.3 REGRESIÓN LINEAL MÚLTIPLE

#### 3.1.3.1 El coeficiente de correlación múltiple

El coeficiente de correlación múltiple,  $\rho_{Y,1\dots,p}$ , es una medida del grado de asociación lineal entre la variable dependiente, Y, y el conjunto de variables  $X_1, \dots, X_p$ . Es la máxima correlación entre Y y cualquier Y' que sea combinación lineal de  $X_1, \dots, X_p$ . Su valor está comprendido entre 0 y 1.

Si es próximo a 1, el ajuste del plano de regresión será casi perfecto, y si es próximo a 0, el plano de regresión no mejorará la predicción de Y sobre la predicción obtenida con la media muestral de Y.

El estimador muestral de  $\rho_{Y.1\dots p}$  es el coeficiente de correlación múltiple muestral, R. En el caso particular de una única variable independiente, el coeficiente de correlación múltiple coincide con el valor absoluto del coeficiente de correlación simple.

Como en el modelo de regresión simple, también en el modelo múltiple la variabilidad total puede descomponerse en la variabilidad explicada por la regresión y la variabilidad residual:

$$SC_{total} = SC_{reg} + SC_{res} \quad (\text{ecuación 3.13})$$

donde  $SC_{total}$ ,  $SC_{reg}$  y  $SC_{res}$  se definen exactamente igual que en el modelo simple. En este caso, el estadístico F proporcionado por la tabla de análisis de la varianza, permite contrastar la hipótesis nula de que la pendiente del plano de regresión es igual a cero, es decir:

$$H_0: \beta_1 = \dots = \beta_p = 0$$

Además se verifica que:

$$R^2 = \frac{SC_{reg}}{SC_{tot}} \quad (\text{ecuación 3.14})$$

donde  $R^2$  es el cuadrado del coeficiente de correlación múltiple muestral, estimación del cuadrado del coeficiente de correlación múltiple  $\rho^2_{Y.1\dots p}$ , al que se denomina coeficiente de determinación. Es decir, el coeficiente de determinación se define a partir de la correlación múltiple y, en el caso particular de una única variable independiente, coincidirá con el cuadrado del coeficiente de correlación simple.

El coeficiente de determinación presenta el inconveniente de que, a mayor número de variables en el modelo, mayor es su valor, por lo que, en general, se considera que  $R^2$  tiende a sobrestimar el verdadero valor de  $\rho^2_{y_1, \dots, p}$ .

El coeficiente de determinación ajustado por el número de observaciones y el número de variables independientes incluidas en la ecuación de regresión:

$$R^2_a = \frac{(n-1)R^2 - p}{n-1-p} \quad (\text{ecuación 3.15})$$

corrige la sobrestimación de  $R^2$ .

### 3.1.3.2 Estadísticos para las variables independientes

Una alternativa para comparar la contribución de las distintas variables al modelo sería comparar los coeficientes correspondientes en la ecuación de regresión: a mayor coeficiente, mayor influencia de los valores de la variable correspondiente en la estimación del valor de la variable dependiente. Sin embargo, los coeficientes de la ecuación de regresión presentan el inconveniente de que su magnitud es relativa.

Teniendo en cuenta que el coeficiente asociado a una variable independiente es igual al incremento (positivo o negativo) que se produciría en la variable dependiente al variar en una unidad el valor de la independiente, un mismo coeficiente para dos variables en distintas unidades de medida podría indicar una importancia relativa muy distinta.

Los coeficientes de regresión tipificados son los coeficientes de las variables cuando la ecuación de regresión se expresa como función de las variables tipificadas. Dado que al tipificar las variables se homogeneiza la unidad de medida, el coeficiente de regresión tipificado se puede interpretar como una medida de la contribución relativa de la variable

correspondiente al modelo. El plano de regresión construido a partir de la tipificación de las variables pasará por el origen y, por tanto, el coeficiente correspondiente al término independiente será igual a cero.

Mediante el estadístico F asociado a la descomposición de la varianza contrastamos la hipótesis nula de que la pendiente del plano de regresión era igual a cero:

$$H_0: \beta_1 = \dots = \beta_p = 0$$

El rechazar esta hipótesis no implica que, aunque el conjunto de las variables independientes mejore la estimación de los valores de la variable dependiente respecto a la media, todas ellas contribuyan a la mejora. Para comprobarlo, contrastaremos la hipótesis anterior sobre cada parámetro en particular:

$$H_0: \beta_i = 0 \quad \forall i = 1, \dots, p$$

En este caso, la hipótesis nula significa que la variable  $X_i$  no mejora la predicción de  $Y$  sobre la regresión obtenida con las  $p-1$  variables restantes. El estadístico de contraste (  $T$  ) es igual a :

$$\frac{B_i}{S_{B_i}} \quad (\text{ecuación 3.16})$$

donde  $S_{B_i}$  es el error típico del coeficiente  $B_i$ . Si el p-valor asociado al estadístico de contraste  $T$  es menor que  $\alpha$ , se rechazará la hipótesis nula al nivel de significación  $\alpha$ .

La hipótesis nula anterior también podría plantearse para el término independiente del modelo:

$$H_0: \beta_0 = 0$$

En este caso, se interpretaría como que el plano de regresión pasa por el origen.

## 3.2 SERIES TEMPORALES

### 3.2.1 CONCEPTOS TEÓRICOS

Una serie temporal se puede considerar como un conjunto de observaciones de una variable tomadas en intervalos regulares de tiempo. El análisis de series temporales estudia cómo construir un modelo para explicar la estructura y prever la evolución de una serie. En particular, la metodología Box-Jenkins es un procedimiento de análisis estadístico para ajustar a una serie un tipo especial de modelos denominados ARIMA (Autorregresive Integrated Moving Average). Dicho análisis consta de las siguientes fases:

- Identificación de la estructura del modelo: decidir que transformación aplicar a la serie observada para convertir el proceso subyacente en estacionario y determinar la estructura del proceso estacionario resultante.
- Estimación de los parámetros del modelo para el proceso estacionario.
- Diagnósis del modelo: Comprobar si se satisface la hipótesis de que los residuos obtenidos al ajustar el modelo son una realización de un proceso de ruido blanco.

#### 3.2.1.1 Procesos estacionarios

El marco teórico del análisis de series temporales es el de los procesos estocásticos. Si se define un proceso estocástico como una sucesión de variables aleatorias  $\{X_t\}$ ,  $t=1, \dots, n$ , ordenadas en el tiempo, una serie temporal puede ser considerada como una realización o trayectoria de un proceso estocástico. Es decir, el valor observado de la serie en el instante  $t$  puede ser considerado como una muestra aleatoria de tamaño uno de la variable  $X_t$  del proceso estocástico definida en dicho instante.

Diremos que  $X_t$  y  $X_{t'}$  están separadas por  $k$  retardos si  $|t - t'| = k$ .

3.2.1.2 Función de autocorrelación. Se denomina función de autocorrelación de un proceso, a la función que describe las correlaciones en dos variables cualesquiera del proceso:

$$\rho(t, k) = \text{Corr}(X_t, X_{t-k}) \quad \forall t, k \quad (\text{ecuación 3.17})$$

3.2.1.3 Proceso estacionario. Un proceso estocástico  $\{X_t\}$  es estacionario (en covarianzas) si:

- La media de  $X_t$  es constante:  $E(X_t) = \mu_t = \mu \quad \forall t.$
- La varianza de  $X_t$  es constante:  $V(X_t) = \sigma_t^2 = \sigma^2 \quad \forall t.$
- La correlación entre  $X_t$  y  $X_{t+k}$  depende únicamente del número de retardos que las separan:

$$\rho(t, k) = \rho_k \quad \forall t, \quad \forall k = 0, 1, 2, \dots \quad (\text{ecuación 3.18})$$

3.2.1.4 Autocorrelación simple de orden  $k$ ,  $\rho_k$ . Dado un proceso estacionario, la autocorrelación parcial de orden  $k$  es la correlación entre variables separadas por  $k$  retardos. La representación de los coeficientes de autocorrelación en función del retardo se denomina función simple (fas).

3.2.1.5 Autocorrelación parcial de orden  $k$ ,  $\alpha_k$ . Dado un proceso estacionario, la autocorrelación parcial de orden  $k$  es la correlación parcial entre variables separadas por  $k$  retardos eliminando el efecto de las  $k-1$  variables intermedias. La representación de los coeficientes de autocorrelación parcial en función del retardo se denomina función de autocorrelación parcial (fap).



3.2.1.6 Proceso de ruido blanco. Un proceso estacionario  $\{a_t\}$  es un ruido blanco si:

- $E\{a_t\} = 0 \quad \forall t.$
- $V\{a_t\} = \sigma^2 \quad \forall t.$
- $\text{Corr}(a_t, a_{t-k}) = \rho_k = 0, \quad k = \pm 1, \pm 2, \dots, \forall t.$
- $a_t$  sigue una distribución Normal  $\forall t.$

3.2.1.7 Proceso AR(p). Un proceso  $\{X_t\}$  es autorregresivo de orden p si:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t + c \quad (\text{ecuación 3.19})$$

donde  $\{a_t\}$  es un proceso de ruido blanco con  $a_t$  independiente de  $X_{t-h}$ ,  $\forall h \geq 1, \forall t.$

Si B es el operador retardo:

$$B X_t = X_{t-1}, \dots, B^k X_t = X_{t-k} \quad (\text{ecuación 3.20})$$

la ecuación que resulta es:

$$\phi_p(B) X_t = a_t + c \quad (\text{ecuación 3.21})$$

donde:

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (\text{ecuación 3.22})$$

Un proceso estacionario y autorregresivo AR(p) se caracteriza porque los p primeros coeficientes de la fap son no nulos y el resto cero, y la fas, en general, presenta muchos coeficientes no nulos que decrecen con el retardo con mezcla de exponenciales y sinusoidales.

3.2.1.8 Proceso MA(q). Un proceso  $\{X_t\}$  es una media móvil de orden q si:

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} + c \quad (\text{ecuación 3.23})$$

donde  $\{a_t\}$  es un proceso de ruido blanco. Considerando el operador retardo

B, la ecuación que resulta es:

$$X_t = \theta_q(B) a_t + c \quad (\text{ecuación 3.24})$$

donde:

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (\text{ecuación 3.25})$$

Todo proceso de media móvil es estacionario. Un proceso de media móvil MA(q) se caracteriza porque los q primeros coeficientes de la fas son no nulos y el resto cero, y la fap, en general, presenta muchos coeficientes no nulos que decrecen con el retardo con mezcla de exponenciales y sinusoidales.

3.2.1.9 Proceso ARMA(p,q). Un proceso  $\{X_t\}$  es autorregresivo de medias móviles de ordenes p y q si:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} + c \quad (\text{ecuación 3.26})$$

donde  $\{a_t\}$  es un proceso de ruido blanco con  $a_t$  independiente de las  $X_t$ . Considerando el operador retardo B, la ecuación que resulta es:

$$\phi_p(B) X_t = \theta_q(B) a_t + c \quad (\text{ecuación 3.27})$$

donde:

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (\text{ecuación 3.28})$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (\text{ecuación 3.29})$$

Dado que los procesos ARMA son una mezcla de procesos AR y procesos MA, la fas y la fap de un proceso estacionario ARMA(p,q) será una superposición de las fas y fap de los procesos AR(p) y MA(q) correspondientes.

### 3.2.1.10 Proceso ARIMA(p, d, q).

Un proceso  $\{X_t\}$  es autorregresivo integrado de medias móviles de ordenes  $p$ ,  $d$ , y  $q$  si:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \nabla^d X_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t + c \quad (\text{ecuación 3.30})$$

donde  $\{a_t\}$  es un proceso de ruido blanco y  $\nabla$  es el operador diferencia:

$$\nabla X_t = (1 - B)X_t = X_t - X_{t-1}, \dots, \nabla^d X_t = (1 - B)^d X_t = (1 - B)^{d-1} (X_t - X_{t-1}) \quad (\text{ec. 3.31})$$

$$\text{Equivalentemente:} \quad \phi_p(B) W_t = \theta_q(B) a_t + c \quad (\text{ecuación 3.32})$$

$$\text{donde:} \quad W_t = \nabla^d X_t \quad (\text{ecuación 3.33})$$

Luego el proceso ARIMA(p, d, q) para  $X_t$  es equivalente al proceso ARMA(p, q) para  $W_t$ . Si el proceso  $\{W_t\}$  es estacionario el proceso ARIMA para  $X_t$  quedará caracterizado por la fase y la fap del proceso  $\{W_t\}$  y por  $d$ , el orden de diferenciación del proceso original  $\{X_t\}$ .

3.2.1.11 Proceso Estacional. Un tipo especial de procesos no estacionarios son los estacionales. Un proceso  $\{X_t\}$  será estacional cuando presente una pauta regular de comportamiento periódico. Si la pauta estacional es constante y con periodo  $s$ , en general aplicando el operador:

$$(1 - B^s) X_t = X_t - X_{t-s} \quad (\text{ecuación 3.34})$$

La estacionalidad podrá ser eliminada. En ocasiones será necesaria más de una diferencia estacional para desestacionalizar el proceso:

$$(1 - B^s)^D X_t = (1 - B^s)^{D-1} (1 - B^s) X_t = (1 - B^s)^{D-1} (X_t - X_{t-s}) \quad (\text{ecuación 3.35})$$

siendo  $D$  el orden de diferenciación estacional.

### 3.2.1.11.1 Proceso ARIMA(p,d,q) (P,D,Q)<sub>s</sub>.

Un proceso estacional  $\{X_t\}$ , con periodo  $s$  conocido, es un proceso autorregresivo integrado de medias móviles de ordenes  $p, d, q, P, D$  y  $Q$  si:

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^p X_t = \theta_q(B)\Theta_Q(B^s)a_t + c \quad (\text{ecuación 3.36})$$

donde  $\{a_t\}$  es un proceso de ruido blanco,  $d$  y  $D$  son los ordenes de diferenciación regular y estacional, respectivamente, y

$$\Phi_p(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_p B^{ps} \quad (\text{ecuación 3.37})$$

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (\text{ecuación 3.38})$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (\text{ecuación 3.39})$$

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs} \quad (\text{ecuación 3.40})$$

Equivalentemente:

$$\Phi_p(B^s)\phi_p(B)W_t = \theta_q(B)\Theta_Q(B^s)a_t + c \quad (\text{ecuación 3.41})$$

donde:

$$W_t = (1-B)^d(1-B^s)^p X_t \quad (\text{ecuación 3.42})$$

Luego el proceso ARIMA estacional para  $X_t$  es equivalente al proceso ARMA estacional para  $W_t$ . Si el proceso  $\{W_t\}$  es estacionario, el proceso ARIMA estacional para  $X_t$  quedará caracterizado por la  $fap$  y la  $fap$  del proceso  $\{W_t\}$  y por  $d$  y  $D$ , los ordenes de diferenciación regular y estacional, respectivamente, del proceso original  $\{X_t\}$ .

Las funciones de autocorrelacion simple y parcial de un proceso ARMA estacional son una mezcla de sus componentes regulares y estacionales, presentando la siguiente estructura:

- Los retardos bajos presentan la estructura de la parte regular, mientras que los retardos estacionales ( $s, 2s, \dots$ ) presentan la de la parte estacional.
- En la  $f_{as}$ , alrededor de los retardos estacionales se manifiesta la interacción entre la parte regular y la estacional, con la repetición, a ambos lados de cada retardo estacional, de la  $f_{as}$  de la parte regular.
- En la  $f_{ap}$ , a la derecha de cada coeficiente estacional se repite la estructura de la  $f_{ap}$  correspondiente a la parte regular, y a la izquierda de los coeficientes estacionales, la de la  $f_{as}$  correspondiente a la parte regular.

### 3.2.2 ANÁLISIS DE UNA SERIE TEMPORAL MEDIANTE LA METODOLOGÍA BOX-JENKINS

#### 3.2.2.1 Identificación del modelo ARIMA

El primer paso en el análisis de una serie es la identificación del modelo. Si la serie no presenta estacionalidad, determinar los ordenes del modelo  $ARIMA(p, d, q)$ , y si la serie presenta estacionalidad del periodo  $s$ , los ordenes del modelo  $ARIMA(p, d, q)(P, D, Q)_s$ . Para identificarlos comenzaremos por determinar si la serie es o no estacionaria, es decir, analizaremos si la varianza es o no constante, si la media es o no constante y si la serie presenta o no estacionalidad.

Si la varianza de la serie no fuera constante se trataría de encontrar una transformación de la serie que estabilizara la varianza y, caso de encontrarla, al análisis proseguiría sobre la serie transformada. Si la media de la serie no fuera constante, se diferenciaría regularmente la serie hasta estabilizar la media, identificando de esta forma el parámetro  $d$  (número de veces que hay que diferenciar regularmente la serie para obtener una serie estacionaria en medias).

Si la serie fuera estacional, se tomarían diferencias estacionales hasta eliminar la estacionalidad, identificando de esta forma el parámetro  $D$  (número de diferencias estacionales). Si la serie no es estacionaria en

medias y es estacional, en ocasiones, al diferenciar estacionalmente únicamente, además de eliminar la estacionalidad, se estabiliza la media.

En dicho caso no sería necesario, ni conveniente, considerar la diferenciación regular. Una vez determinados los parámetros  $d$  y  $D$  el siguiente paso será identificar los parámetros  $p$  &  $q$ , y en caso de que la serie presentara estacionalidad, los parámetros  $P$  y  $Q$ .

#### 3.2.2.2 Homocedasticidad.

Las series pueden presentar una variabilidad no constante, es decir, la varianza  $\sigma_t^2$  puede ser dependiente de  $t$ . Una forma de detectar si la varianza es constante es agrupar las observaciones por periodos de tiempo con el mismo número de observaciones (por ejemplo, si los datos corresponden a observaciones mensuales, agruparlos en años) y realizar la prueba de Levene para la homogeneidad de varianzas de la serie observada en los  $K$  grupos definidos. Esta prueba permite comprobar la hipótesis de que los  $K$  grupos proceden de poblaciones con varianza común.

Si la hipótesis de homogeneidad de varianzas fuera rechazada podría ser debido a que la dispersión de la serie cambia con la tendencia central. En concreto, una situación frecuente es que en los periodos en los que la tendencia central es grande también lo sea la dispersión. En esta situación existe una familia de transformaciones que puede, en general, estabilizar la varianza. En la practica se utilizan transformaciones de tipo:

$$T(X_t) = \begin{cases} X_t^p & p \neq 0 \\ \text{Ln}(X_t) & p = 0 \end{cases} \quad (\text{ecuación 3.43})$$

donde  $p$  es el múltiplo de un medio mas próximo al poder de transformación proporcionado por el gráfico de nivel y de dispersión. El gráfico de nivel y dispersión representa, en cada uno de los  $K$  grupos, el logaritmo neperiano del rango intercuartílico de la serie  $X_t$  (en ordenadas) frente al

logaritmo neperiano de la mediana (en abcisas). El poder de transformación es igual a 1 menos la pendiente de la recta de regresión mínimo-cuadrática ajustada a los  $K$  puntos. Esta familia de transformaciones no solo permite estabilizar la varianza sino que, incluso, puede proporcionar normalidad. Por otro lado, sólo está definida para datos positivos, por lo que en ocasiones, antes de realizar la transformación será necesario sumar un misma constante a todos los valores de  $X_t$ .

### 3.2.2.3 Análisis de la estabilidad en medias y de la estacionalidad

Puede suceder que la serie observada presente un patrón regular de comportamiento a largo plazo con tendencia creciente o decreciente. En dicho caso, la serie no sería estable en medias. Si en el gráfico de la serie la presencia de tendencia creciente o decreciente no es evidente, una alternativa es representar la función de autocorrelación simple estimada (fas). La estimación de la autocorrelación simple de orden  $K$  del proceso subyacente se obtiene como la correlación entre la serie observada,  $X_t = (x_1, \dots, x_n)$ , y la serie retardada  $K$  veces,  $Y_t = (y_1, \dots, y_{n-k})$ , donde  $y_1 = x_{k+1}, \dots, y_{n-k} = x_n$

Si la serie observada presentara tendencia creciente o decreciente su fas representaría estructura positiva con decrecimiento lento hacia cero. La tendencia puede estabilizarse aplicando sucesivas diferencias regulares:

$$W_t = \nabla^d X_t = (1-B)^d X_t \quad (\text{ecuación 3.44})$$

hasta que la serie obtenida,  $W_t$  presente media constante.

Un caso particular de no estacionariedad presente en muchas series es la estacionalidad, entendiendo por estacionalidad una pauta regular de comportamiento periódico en la serie. Si en el gráfico de la serie la presencia de estacionalidad no es evidente, una alternativa es representar la fas estimada. Si la serie observada presentara estacionalidad del periodo  $s$ , su fas presentaría coeficientes altos con decrecimiento lento en los retardos  $s, 2s, 3s \dots$ . La estacionalidad puede eliminarse aplicando

sucesivas diferencias estacionales de periodo  $s$ :

$$W_t = \nabla_s^D X_t = (1 - B^s)^D X_t \quad (\text{ecuación 3.45})$$

hasta que la serie obtenida,  $W_t$ , no presente estacionalidad.

#### 3.2.2.4 Determinación de los órdenes $p$ , $d$ , $q$ , $P$ , $D$ , y $Q$ en el modelo.

Los órdenes  $d$  y  $D$  son el número de veces que hay que diferenciar  $X_t$  de forma regular y de forma estacional, respectivamente, hasta obtener un proceso estacionario. Los órdenes  $p$ ,  $q$ ,  $P$  y  $Q$  se determinarán teniendo en cuenta que, si:

$$W_t = (1 - B)^d (1 - B^s)^D X_t \quad (\text{ecuación 3.46})$$

es el proceso estacionario obtenido al diferenciar  $X_t$   $d$  veces de forma regular y  $D$  veces de forma estacional, entonces  $W_t$  sigue un proceso ARMA estacional:

$$\Phi_p(B^s) \phi_p(B) W_t = \theta_q(B) \Theta_Q(B^s) a_t + c \quad (\text{ecuación 3.47})$$

En otras palabras, los órdenes  $p$ ,  $q$ ,  $P$  y  $Q$ , se determinarán a partir de la estructura de las funciones de autocorrelación simple y parcial de la serie  $W_t$ . La estimación de la autocorrelación parcial de orden  $k$  del proceso subyacente se obtiene como la correlación entre la serie observada,  $X_t = (x_1, \dots, x_n)$ , y la serie retardada  $K$  veces,  $Y_t = (y_1 = x_{k+1}, \dots, y_{n-k} = x_n)$ , eliminando el efecto de las series retardadas 1, 2, ...,  $k-1$  veces.

Para determinar cuál es el proceso del que la serie observada  $X_t$  puede ser considerada como una realización o trayectoria será necesario analizar las estimaciones de las funciones de autocorrelación simple y parcial. En la practica:

- Si la serie observada presenta media constante y no es estacional, los parámetros  $d$ ,  $D$ ,  $P$  y  $Q$  son iguales a 0 y, en consecuencia, el modelo que se



ajustará será un  $ARMA(p,q)$ . Los parámetros  $p$  y  $q$  se estimarán a partir de las funciones de autocorrelación simple y parcial de la propia serie observada. En particular, si los primeros coeficientes de la fap son no nulos y el resto cero, y en la fas los coeficientes decrecen con el retardo con mezcla de exponenciales y sinusoidales, el modelo identificado será un modelo autorregresivo,  $AR(p)$ , de orden  $p$  igual al número de coeficientes no nulos en la fap. Si por el contrario, es la fas la que presenta los primeros coeficientes no nulos y el resto cero, y en la fap los coeficientes decrecen con el retardo con mezcla de exponenciales y sinusoidales, el modelo identificado será un modelo de medias móviles,  $MA(q)$ , de orden  $q$  igual al número de coeficientes no nulos en la fas. Si la fas y la fap son una superposición de las fas y fap descritas en las dos situaciones anteriores, tanto  $p$  como  $q$  serán distintos de cero y su identificación se hará difícil en la práctica, por lo que se recomienda ajustar, inicialmente, un modelo  $ARMA(1,1)$ . Si previamente al análisis de la tendencia y de la estacionalidad hubiera sido necesario considerar una transformación de la serie para estabilizar la varianza,  $p$  y  $q$  se identificarían a partir de las funciones de autocorrelación simple y parcial para la serie transformada.

- Si la serie observada presenta tendencia creciente o decreciente pero no estacionalidad, los parámetros  $D$ ,  $P$  y  $Q$  son iguales a 0 y el parámetro  $d$  es igual al número de veces que ha sido necesario diferenciar regularmente la serie para estabilizar la tendencia. En consecuencia, el modelo que se ajustará será  $ARIMA(p,d,q)$ . Los parámetros  $p$  y  $q$  se estimarán igual que en el apartado anterior, pero a partir de las funciones de autocorrelación simple y parcial de la serie diferenciada.
  
- Si la serie observada presenta estacionalidad del periodo  $s$ , el modelo que se ajustará será un  $ARIMA(p,d,q)(P,D,Q)s$ , donde  $d$  y  $D$  son iguales al número de veces que ha sido necesario diferenciar regular y estacionalmente la serie para estabilizar la tendencia y eliminar la estacionalidad (si la serie original es estable en medias, entonces  $d$  será igual a 0). Los parámetros  $p$ ,  $q$ ,  $P$  y  $Q$  se estimarán a partir de las funciones de autocorrelación simple y parcial de la serie diferenciada regular y estacionalmente. En particular  $p$  y  $q$  se estimarán igual que en los apartados anteriores, pero considerando únicamente los coeficientes

correspondientes a los retardos anteriores al primer retardo estacional. P y Q se estimarán, de la misma forma que p y q, pero considerando exclusivamente los retardos estacionales.

### 3.2.2.5 Ajuste del modelo ARIMA

#### 3.2.2.5.1 *Estimación de los parámetros*

Una vez determinados los ordenes p, d, q, P, D y Q de un proceso ARIMA(p,d,q) (P,D,Q)s con periodo conocido s se trata de estimar los parámetros del modelo a partir de una realización del proceso u observación de la serie  $(x_1, \dots, x_n)$ :

$$(1-\phi_1 B - \dots - \phi_p B^p) (1-\Phi_1 B^s - \dots - \Phi_P B^{Ps}) (1-B)^d (1-B^s)^D X_t = (1-\theta_1 B - \dots - \theta_q B^q) (1-\Theta_1 B^s - \dots - \Theta_Q B^{Qs}) a_t + c \quad (\text{ecuación 3.48})$$

Es decir, se trata de estimar los parámetros  $\beta_i$ ,  $i=0, \dots, p+P+q+Q$ , donde:

$$\beta_i = \phi_i \quad i = 1, \dots, p. \quad (\text{ecuación 3.49})$$

$$\beta_i = \Phi_{i-p} \quad i = p + 1, \dots, P + p. \quad (\text{ecuación 3.50})$$

$$\beta_i = \theta_{i-(p+P)} \quad i = p + P, \dots, p + P + q. \quad (\text{ecuación 3.51})$$

$$\beta_i = \Theta_{i-(p+P+q)} \quad i = p + P + q + 1, \dots, p + P + q + Q. \quad (\text{ecuación 3.52})$$

$$\beta_i = c \quad i = 0. \quad (\text{ecuación 3.53})$$

Si  $B_i$  es la estimación del parámetro  $\beta_i$ , la primera etapa en la validación del modelo consistirá en comprobar si los coeficientes  $B_i$  son significativamente distintos de cero. Para ello, sobre cada parámetro, se planteará la hipótesis nula:

$$H_0: \beta_i = 0$$

Dicha hipótesis puede ser interpretada como que la variable asociada al parámetro  $\beta_1$  no mejora el ajuste con respecto al obtenido con las restantes variables incluidas en el modelo. Si el valor-p asociado al valor estadístico de contraste  $t$  es menor que  $\alpha$ , se rechazará la hipótesis nula al nivel de significación  $\alpha$ .

#### 3.2.2.5.2 *Sobreajuste del modelo*

Para comprobar si algún otro modelo se ajusta mejor conviene utilizar la técnica del sobreajuste. Éste consiste en estimar cada uno de los modelos obtenidos al aumentar en una unidad cada uno de los ordenes  $p$ ,  $q$ ,  $P$  y  $Q$  en el modelo original.

El hecho de que todos los coeficientes correspondientes a un sobreajuste sean significativamente distintos de cero no significa que el modelo correspondiente sea mejor que el original. Para comprobar si es así sería necesario comparar sus residuos. Si al comparar aquellos modelos tales que las estimaciones de todos los parámetros son significativamente distintas de cero, el mejor corresponde a uno de los sobreajustes, se procederá como con el modelo original: se sobreajustará y se comparará con los modelos sobreajustados.

El proceso continuará hasta que ningún modelo mejore el ajuste del modelo de partida. En cualquier caso, se recomienda que los ordenes  $p$ ,  $q$ ,  $P$  y  $Q$  no alcancen valores superiores a 2. Si un modelo admite algún orden superior a 2, puede ser un indicio de que el modelo originalmente seleccionado no es el más adecuado.

Esta situación puede darse como consecuencia de que un modelo autorregresivo de orden grande puede ser expresado, como un modelo de medias móviles de orden pequeño, y viceversa.

### 3.2.2.5.3 Validación del modelo

La validación del modelo consistirá en comprobar si se satisface la hipótesis del modelo ARIMA relativa a que  $\{a_t\}$  es un proceso de ruido blanco. El inconveniente radica en que, dado que los coeficientes del modelo son estimaciones, a partir de la muestra observada, de los verdaderos parámetros, los verdaderos errores  $a_t$  son desconocidos. Por consiguiente, la comprobación de la hipótesis de proceso de ruido blanco se realizará sobre una estimación de los mismos. Si  $\hat{x}_t$  es la estimación de  $x_t$  mediante el modelo ajustado, una estimación de la realización del proceso de ruido blanco  $\{a_t\}$  será el residuo:

$$e_t = \hat{a}_t = x_t - \hat{x}_t \quad (\text{ecuación 3.54})$$

La validación consistirá en probar que la serie residual tenga media cero, que es estable en varianzas, que sus observaciones están incorreladas y que procede de una distribución normal.

### 3.2.2.5.4 Etapas en la construcción de un modelo ARIMA

- Inspeccionar, mediante el gráfico de la serie, la varianza, la media y la estacionalidad.
- Analizar si la serie es estable en varianzas. En caso contrario, buscar una transformación que la estabilice.
- Analizar si la serie es estable en medias. En caso contrario, diferenciar regularmente hasta que la serie obtenida presente media constante.
- Analizar si la serie es estacional. En dicho caso, diferenciar estacionalmente hasta que la serie obtenida no presente estacionalidad.
- Si la serie no es estable en medias y es estacional, probar si diferenciando solo estacionalmente, además de eliminar la estacionalidad, se estabiliza la media.

- Estimar los ordenes de las partes autorregresivas y de medias móviles regulares. Si la serie es estacional, estimar los ordenes de las partes autorregresivas y de medias móviles estacionales.
- Ajustar el modelo elegido y los obtenidos al sobreajustar los ordenes del elegido.
- Seleccionar aquellos modelos tales que las estimaciones de los parámetros sean significativamente distintas de cero y después de compararlos, elegir el mejor modelo. Si el mejor corresponde a uno de los sobrajustes, considerarlo como modelo elegido, sobreajustarlo y compararlo con los modelos sobreajustados. Repetir la misma operación hasta que ningún modelo mejore el ajuste del modelo elegido.
- Validar el modelo elegido. Si los residuos no corresponden a un ruido blanco tratar de mejorarlo.
- Evaluar la capacidad predictiva del modelo elegido considerando un periodo histórico y un periodo de validación. Si el modelo es adecuado en este sentido, obtener predicciones para el futuro con el modelo ajustado.

## CAPÍTULO IV

### LA CUENCA DEL VALLE DE MÉXICO

#### 4.1 BREVE RESEÑA HISTÓRICA

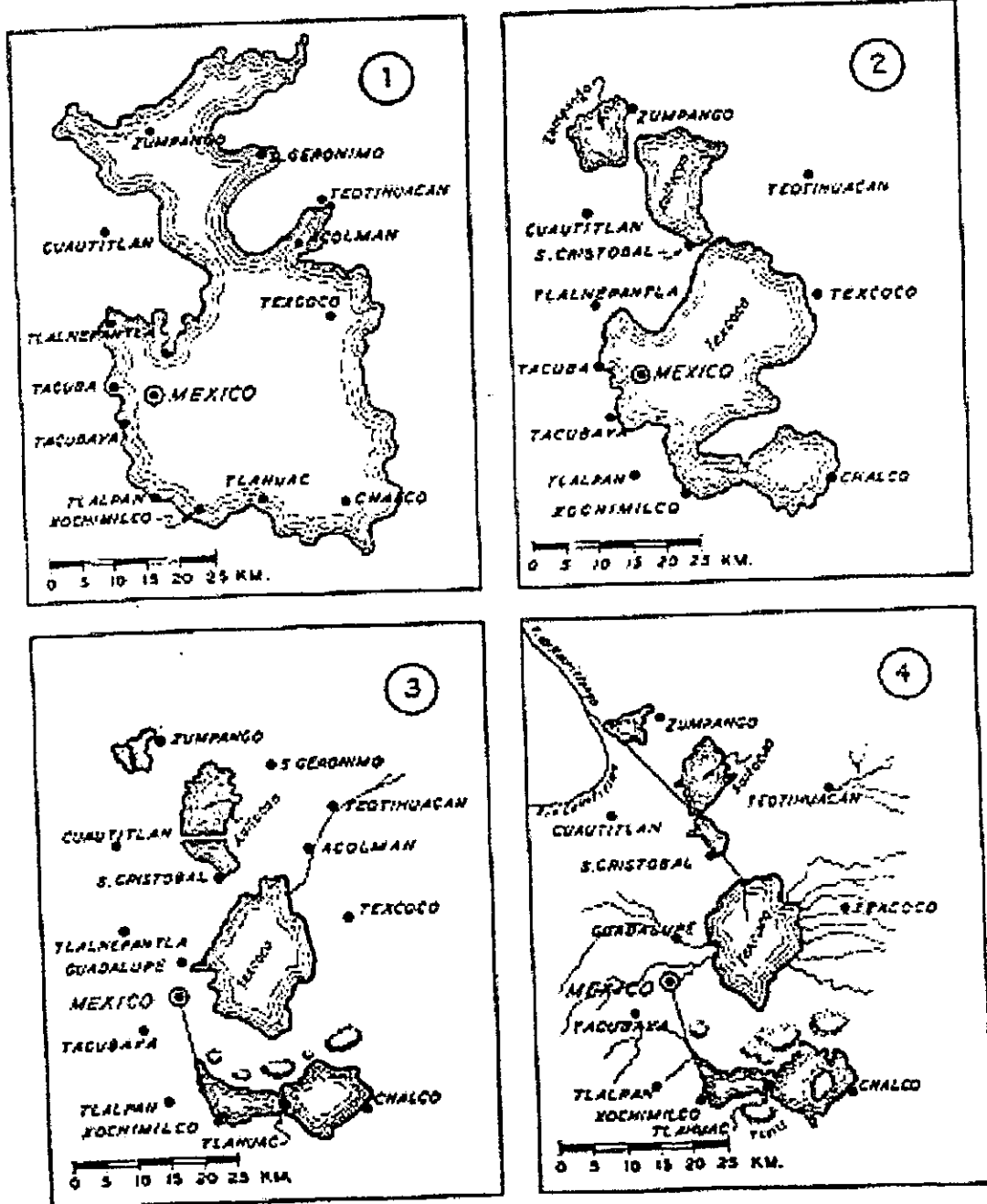
##### 4.1.1 LA ÉPOCA PREHISPÁNICA

Cuando los Mexicas llegaron al valle de México, este era un lago muy grande. En la temporada de secas, parte del lago se evaporaba y parte se resumía, destacándose el centro y en el norte se formaba la laguna conocida después como Ecatepec o San Cristóbal un poco más alta que la de México. En la temporada de lluvias todos los lagos formaban uno solo, crecían sobre sus riberas y cubrían una extensión de casi los 2000 Km<sup>2</sup>; o sea una cuarta parte de la cuenca; el resto lo formaban las laderas cubiertas de grandes árboles, densa vegetación y animales salvajes y también había lugar para los primeros pobladores del valle.

Tenochtitlán se fundó en 1325; antes de la mitad del siglo XV ya se había construido la primera calzada, la de Tacuba que unía el islote con la tierra firme, después se construyeron las calzadas de Tlaltelolco a Azcapotzalco, la de Xochimilco y Coyoacán ordenadas por Itzcoatl. Con esta última y las de Tacuba y Azcapotzalco se formó el primer reducto de las aguas dulces al Suroeste y Oeste de Tenochtitlán, alimentada por los ríos de Churubusco, Mixcoac, Tacubaya y demás corrientes formadas por los manantiales del Desierto de los Leones, Santa Fe y Cuajimalpa, que permitieron una calidad de agua que incrementó la pesca y la industria derivada de ella.

Al comenzar el año 1449, la temporada de lluvias fue muy copiosa, lo que provocó una gran inundación, por la que Netzahualcóyotl aconsejó a Moctezuma I construir diques o muros de madera, piedra y arcilla, el dique del cerro de la Estrella con longitud de 16 km., de Iztapalapa a Atzacolco, en dirección sur norte conocido después como el Albarradón de Netzahualcóyotl; a

Mapa 1  
RETROCESO DE LOS LAGOS EN EL VALLE DE MEXICO



- 1) Los límites aproximados durante la época diluvial.
- 2) A comienzos del siglo XVI.
- 3) A comienzos del siglo XIX.
- 4) En el año 1889.

Plano 2. Evolución de los lagos en el valle de México. (Según Elizabeth Schilling. "Die schwimmenden Gärten con Xochimilco". Kiel, 1938, fig.3).

Chalco y Xochimilco y el de Mexicaltzingo, con estas obras dotadas de compuertas y esclusas se separaron completamente las aguas dulces de las saladas y se controlaron las aguas de la laguna de Texcoco y las de los lagos del sur; después hubo otra inundación mayor, para la que fue necesario reconstruir Tenochtitlán; al final de la 2a. década del s. XVI llegaron los españoles... La capital mexicana cayó el 13 de agosto de 1521.

#### 4.1.2 LA COLONIA

Los españoles no entendieron el papel regulador de los canales y acequias, los destruyeron, conservando algunos de los principales canales. A esto siguió una tala indiscriminada de los montes cercanos, que con el tiempo, las consecuencias se presentaron en grandes y seguidas inundaciones de la ciudad. En esta época hubo varios proyectos para salvar la ciudad, dos de los más importantes fueron creados por Enrico Martínez, el primero era muy completo, pues abarcaba el desagüe y control de todas las lagunas, solamente se llevó a cabo el segundo que se refería a la laguna de Zumpango y al río Cuautitlán que fueron desaguados por el túnel o socavón de Huehuetoca y así la cuenca tuvo ya una salida, por lo incompleto del trabajo (por razones económicas) el túnel se derrumbó y obstruyó la salida provocando una inundación y la evacuación de gran parte de la población de la ciudad.

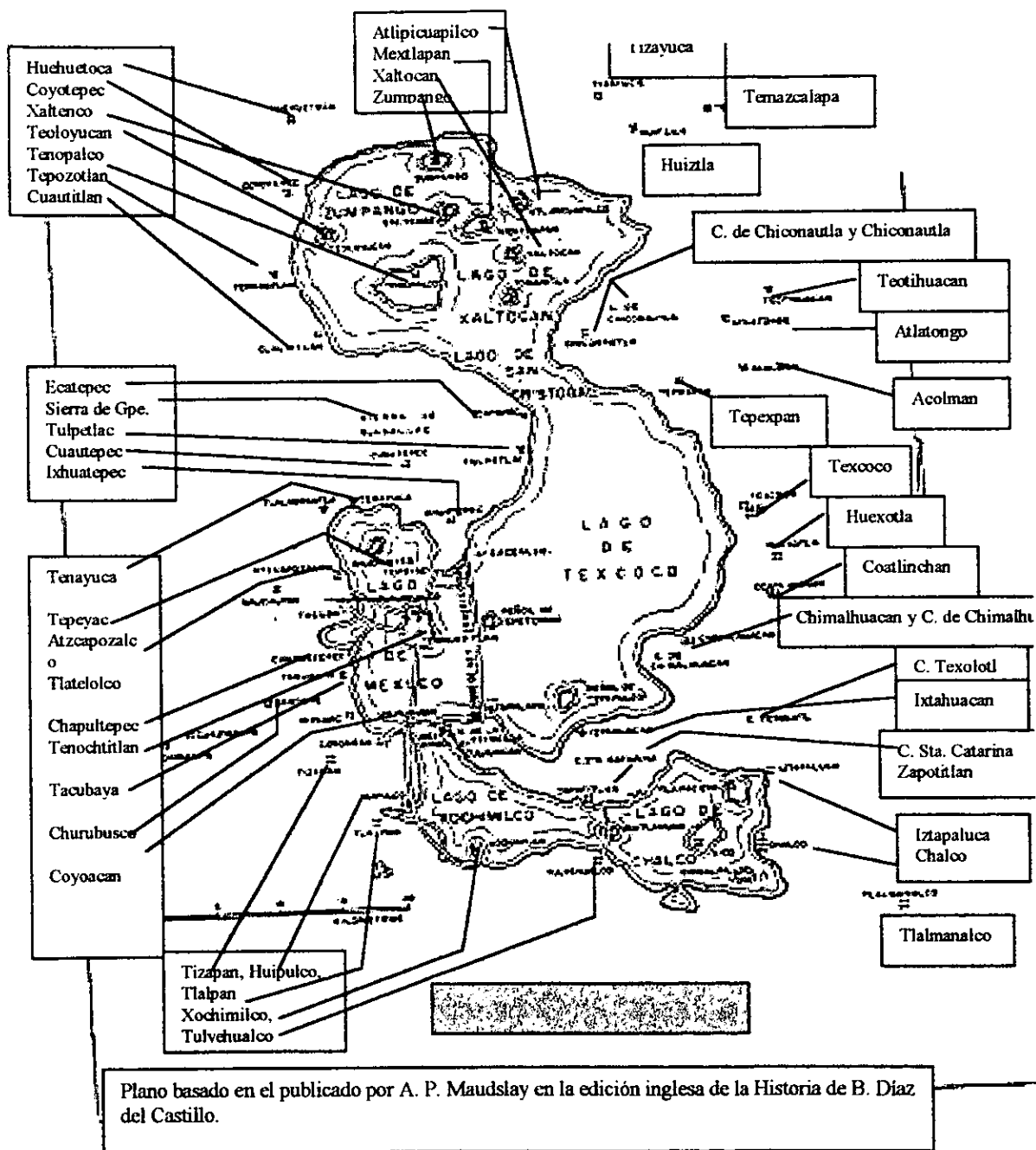
Siguieron las inundaciones y los proyectos, uno de estos contenía un plan para el desagüe general que incluía un túnel para Tequixquiac, pero no se realizó; en 1788 el Real Tribunal del Consulado de México dió por terminado el Tajo de Nochistongo.

#### 4.1.3 EN EL MÉXICO INDEPENDIENTE

Durante la guerra de independencia se paralizaron las obras. Ya en 1856 se aprobó el proyecto del ingeniero de De Garay, en el que se eligió a Tequixquiac para el canal y túnel y trazando varios canales para introducir las aguas de Chalco y Xochimilco a las atarjeas de la ciudad, comunicando el lago de Xochimilco con el tajo de Nochistongo y el de Chalco con el de Zumpango con lo que se dominaban las aguas y se favorecían los transportes y el riego de tierras, pero por el periodo de sequía, la guerra civil y la



Mapa 2  
LAGOS DEL VALLE DE MEXICO  
A FINES DEL SIGLO XV Y PRINCIPIOS DEL XVI



Plano 3. Los lagos a finales del s. XV y principios del XVI.

Fuente: SRH. Ingeniería Hidráulica en México. Vol. XIV núm. 3, 1960, 130p.

intervención francesa no se realizó el proyecto además de que no se pagó el precio al autor por el proyecto.

En la **República**, el gobierno continuaba y paraba las obras. En la **Dictadura** se terminaron los trabajos de De Garay, terminando el canal de desagüe desde San Lázaro y dando salida por el túnel de Tequixquiac en 1900.

- En 1911 Comenzó la desecación del lago de Texcoco.
- En 1915 desapareció el último resto de la laguna de México, con la desecación de la laguneta existente en la Lagunilla.
- En 1925 se inició el alcantarillado de la ciudad.
- En 1937 se inició el control de las corrientes del poniente por medio de presas, canales y túneles para desaguarlos a los ríos de Tlalnepantla.
- En 1940 se comenzó la construcción del segundo túnel de Tequixquiac, inaugurado en 1946.

#### 4.2 CONFORMACIÓN ACTUAL DE LA CUENCA DEL VALLE DE MÉXICO

La República Mexicana esta formada por 37 Regiones Hidrológicas y en la Región Hidrológica No. 26 "Río Pánuco", se ubica la cuenca del valle de México, situada en el extremo sur del altiplano, tiene una superficie de 9600 Km<sup>2</sup>, de la cual alrededor del 30% es plana y situada a una altura media de 2,240 m.s.n.m.

La precipitación media anual es del orden de los 705 mm. concentrados en cuatro meses del año, de los cuales el 80% se evapotranspira, el 11% se infiltra en el subsuelo y del 9% restante, una pequeña parte es regulada para su aprovechamiento y al resto se le da salida por el sistema general de drenaje para evitar inundaciones.

Dentro de la cuenca y en la parte mas baja del valle, se localiza la Zona Metropolitana de la Ciudad de México, en la actualidad la mancha urbana se ha extendido aproximadamente 1,300 Km<sup>2</sup>, ocupados por el Distrito Federal y 27 municipios conurbados del Estado de México; urbanización que ha reducido las zonas de recarga natural del acuífero, alterando el funcionamiento de los cauces provocando la torrencialidad de las corrientes y con ello las

m<sup>3</sup>/seg, los gobiernos de los estado de México e Hidalgo 15.53 y 1.93 respectivamente, 2.34 son explotados por particulares para usos industriales y 4.93 m<sup>3</sup>/seg para uso agropecuario.

En el Distrito Federal el 98% de la población cuenta con el servicio de agua potable mediante tomas domiciliarias y el 2% restante se abastece por medio de carros tanque.

En los municipios conurbados que conforman la Zona Metropolitana de la Ciudad de México la cobertura alcanza el 90%, el resto se abastece por medio de hidrantes públicos y carros tanque.

#### 4.3 SISTEMA HIDROLÓGICO DE LA CUENCA DEL VALLE DE MÉXICO

Siendo el valle de México una cuenca cerrada, a través del tiempo se han llevado a cabo obras para controlar y desfogar los volúmenes que se generan básicamente en época de lluvias, al efecto se han construido estructuras como el Tajo de Nochistongo con los túneles de Tequixquiac-Gran Canal del Desagüe y actualmente el moderno Sistema de Drenaje Profundo de la Ciudad de México.

El sistema hidrológico del valle de México, está constituido por 28 presas, túneles, canales, vasos reguladores, así como por 40 ríos de importancia por su caudal y extensión.

El Sistema se divide en 3 áreas principales:

4.3.1 RÍOS DEL ORIENTE, que comprende desde el río de las Avenidas de Pachuca al norte de la cuenca, hasta el río de la Compañía al sur, teniendo su afluente denominado "Papalote" y cuyos caudales son regulados en la presa "El Manantial" ubicado en las inmediaciones de Tizayuca, Hidalgo. Los ríos San Juan, Papalotla, Xalapango, Coxcacoco, Texcoco, Chapingo, San Bernardino, Santa Mónica, Coaxtepec, de la Compañía con su afluente el San Francisco, todos ellos desembocan al Lago de Texcoco.

4.3.2 RÍOS DEL SUR, que desembocan también al Lago de Texcoco, al poniente se localiza el río Churubusco, el cual a través del Canal Nacional recibe las aportaciones del río Buenaventura y su afluente el San Juan de Dios; Los ríos

Santiago, San Gregorio y San Lucas, desembocan en el lago de Xochimilco terminando los ríos del Sur al Oriente con el río Amecameca.

4.3.3 EN LA PARTE PONIENTE, del valle de México se localizan las corrientes de mayor importancia en el aspecto de control de avenidas y en cierta forma de peligro, ya que cruzan diferentes áreas densamente pobladas; las principales corrientes en esta parte son 26, las cuales forman los siguientes sistemas:

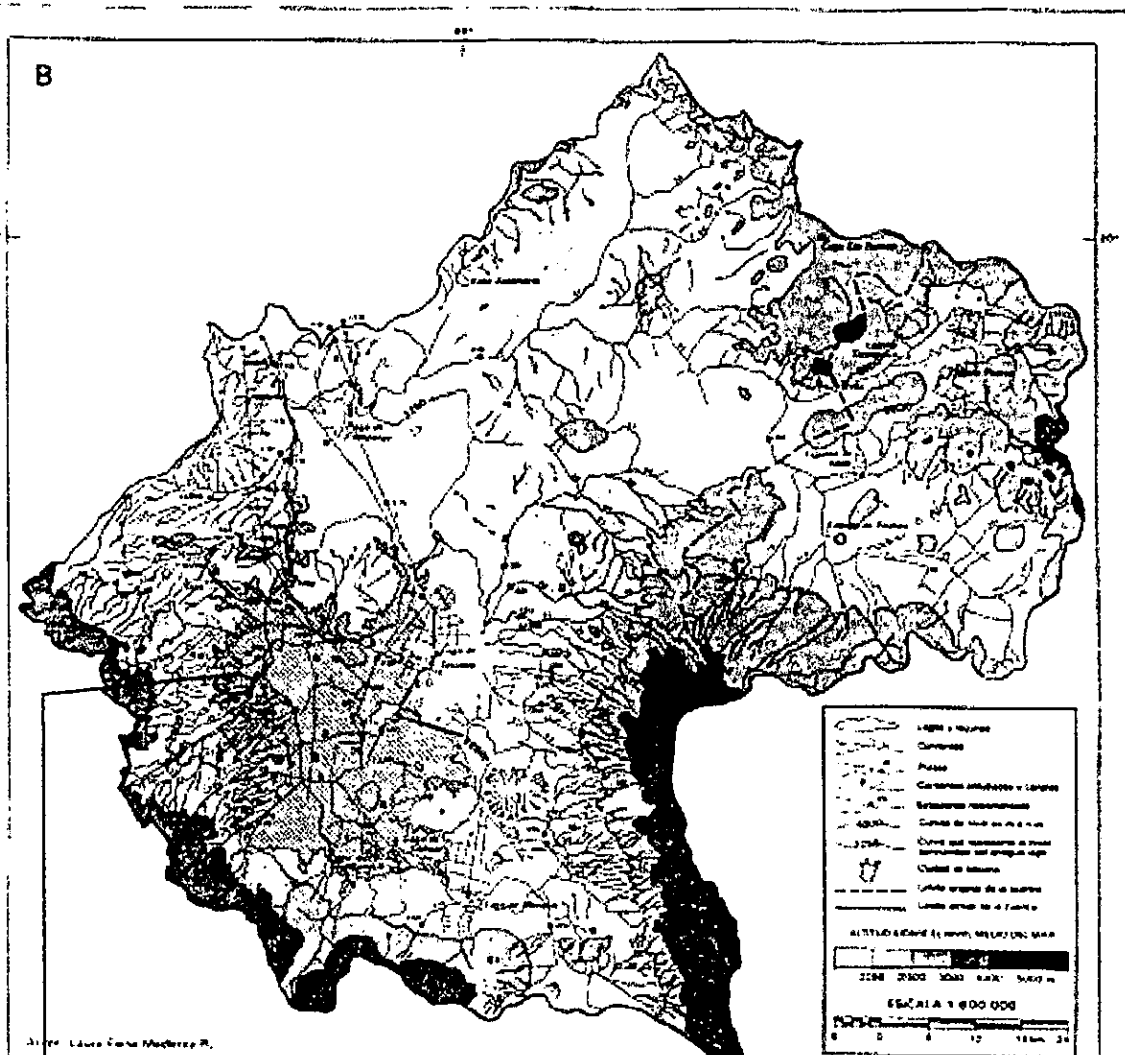
**Sistema de presas Anzaldo-Texcalatlaco**, que incluye el río Eslava, Magdalena, Contreras, Anzaldo, San Jerónimo, Coyotes Texcalatlaco. A su vez este sistema se encuentra conectado al interceptor poniente.

**Sistema de Presas**, comprende el control de avenidas de las siguientes corrientes: Mixcoac, Becerra, Tacubaya, Dolores, Tecamachalco, El Ocote, San Joaquín y el Tornillo, las que son controladas en este sistema y drenadas a través del canal del Tornillo o del Interceptor Poniente hasta el río de los Remedios, que a su vez las conduce al vaso de Cristo.

**Sistema Hondo-Sordo**, este sistema se inicia en la presa derivadora del río Hondo la cual a través de un túnel vierte sus aguas en el río sordo, mismas que una vez controladas a través de la presa que lleva el nombre de El Sordo son descargadas en el río de los Remedios; a esta misma corriente son descargadas las aguas de los ríos "Los Cuartos" y "Totolica", ya controlados por ambas presas de los mismos nombres.

**Sistema Río Chico de los Remedios**, este sistema se compone de las presas: Las Julianas, los Arcos, el Colorado y la Colorada, obras de control de las aguas de los ríos Chico de los Remedios, El Colorado y San Mateo Nopala; en esta última corriente desemboca el túnel que viene de la barranca de Tepetlaco, todas estas aguas confluyen hacia el río Chico de los Remedios el cual es afluente del vaso del Cristo al igual que el río de los Remedios.

Los cuatro sistemas mencionados controlan las avenidas de las aguas pluviales que se precipitan en la serranía del poniente que comprende desde el Ajusco hasta San Mateo Nopala.



El círculo en el mapa muestra la ubicación de la estación Totolica.

En el anexo uno se muestran los nombres de las corrientes, estaciones hidrológicas, presas y corrientes entubadas que están dentro de la cuenca del valle de México.

#### 4.4 ESTACIÓN HIDROMÉTRICA TOTOLICA

NOMBRE DE LA ESTACIÓN: TOTOLICA

##### CARACTERÍSTICAS HIDROGRÁFICAS:

Zona Hidrológica: III (Ciudad de México)  
 Colector General: Río Hondo o de los Remedios  
 Afluente: Río Totolica  
 Corriente: Río Totolica  
 Área de la Cuenca: 23.5 Km<sup>2</sup>

CORRIENTE: RÍO TOTOLICA

Se genera en las laderas de la sierra las Cruces, en una zona topográfica accidentada y cubierta vegetal boscosa. Su dirección inicial es al Oriente la que paulatinamente cambia hacia el Noreste, para descargar sus aguas en el río Hondo. Sobre su cauce se construyó la presa Totolica para control de avenidas; y la Derivadora chica Atoto, para el aprovechamiento de sus caudales. La longitud de esta corriente hasta el sitio de la estación Totolica es de 15.2 km. y su pendiente media de 0.0428. Su principal afluente es el arroyo Cruz Blanca, al que recibe por la margen derecha. Los manantiales que se encuentran en su cuenca son los denominados Cruz Blanca y Totolica.

##### 4.4.1 OBJETIVOS DE SU INSTALACIÓN:

Inicialmente se instaló para determinar el régimen hidráulico del río Totolica; pero en la actualidad su función es cuantificar las salidas de la presa Totolica, construida en 1961.

#### 4.4.2 CARACTERÍSTICAS GEOGRÁFICAS:

Coordenadas: Latitud N: 19° 27'55"

Longitud W.G.: 99° 14'40"

##### Ubicación:

La estación se encuentra 4.0 km. aguas abajo de la cortina de la presa Totolica, así como 1.5 km. al oeste del centro de la población de San Bartolo Naucalpan en el municipio del mismo nombre, en el estado de México.

#### 4.4.3 DATOS DE INSTRUMENTACIÓN HIDROMÉTRICA:

##### Estructura para Aforos:

La estructura para aforos consiste en una pasarela de concreto sobre vigas de hierro, provista de barandal. Tiene 0.80 m de ancho y 10.90 m. de longitud.

##### Escala:

Los niveles de la corriente se observan en una escala de mampostería instalada en la margen derecha, la que guarda una posición inclinada y se localiza 1.50 m aguas arriba de la sección de aforos. Su capacidad máxima de registro es de 3.36 m y el cero de la escala esta en la elevación 2273.244 m.s.n.m.

##### Limnógrafo:

Este sitio dispone de limnógrafo para la observación automática de los niveles del río Totolica. La estructura se encuentra 3.00 m aguas arriba de la sección de aforos y consiste en una caseta de tabique y techo de lamina de asbesto, de 2.15 x 2.34 m de área y 3.00 m de altura, donde se alojó el limnógrafo "Stevens".

Complementa a la estructura una galería abierta de 0.70 x 3.72 m en su porción superior y longitud de 3.90 m.

##### Sólidos en suspensión:

No se practica muestreo de sólidos en suspensión en este sitio de aforos.

#### 4.4.4 METODOLOGÍA EMPLEADA EN EL CÁLCULO HIDROMÉTRICO POR LA GERENCIA DE AGUAS DEL VALLE DE MEXICO:

Resumen del periodo de 1987 a 1994.

Durante este periodo el cálculo de volúmenes y gastos escurridos se realizó generalmente entre los meses de junio a noviembre, el cálculo se hizo por interpolación lineal practicado con molinete hidráulico entre 27 y 100 aforos, aunque en algunos años se usaron otros métodos como el de sección mediante sondeos de 0.50 m y el de velocidad con molinete hidráulico.

También se menciona que durante todo el año de 1994 los cálculos de gasto escurrido y volúmenes se hicieron por interpolación lineal y curva de gastos que constaron de 6 aforos, al igual que en 1993, la curva de gastos se hizo en base a 95 aforos y en 1987 se realizaron 2 curvas de gasto con base de 27 aforos. Además se dispuso de las gráficas continuas de un limnigrafo, comparadas con las lecturas de escala diarias efectuadas a las 6, 12, y 18 horas durante todo el periodo. También se puede apreciar que en los meses de enero a mayo y de noviembre a diciembre no se registró escurrimiento en el río Totollica durante el periodo.

#### NOTA:

Este sitio de aforos es operado por la Delegación en el estado y valle de México, en coordinación con la Dirección de Aguas Superficiales de la SARH.



## CAPÍTULO V

### GENERACIÓN DE MODELOS EN LOS DATOS DE GASTOS MEDIOS DIARIOS DE LA ESTACIÓN HIDROMÉTRICA TOTOLICA

#### 5.1 INTRODUCCIÓN

Los recursos hidráulicos en México se encuentran distribuidos desigualmente a lo largo del país, es decir, algunas zonas carecen de agua y otras son abundantes en la misma. El centro y sur de la república se encuentran densamente poblados, además de que existen zonas industriales y de riego, y la creciente concentración poblacional en zonas de recarga natural ha provocado insuficiencia en el abastecimiento de agua por lo que se requiere de presas. Se escogió la estación hidrométrica Totolica que en la actualidad cuantifica la salida de agua de la presa del mismo nombre.

#### 5.2 DATOS DE ENTRADA

Los datos hidrológicos se obtuvieron del "Boletín hidrológico datos del Valle de México", de los números 40 al 47, equivalentes al periodo anual de 1987-1994, se escogieron estos boletines por estar disponibles, los siguientes al año 94, estaban en proceso. Estos boletines se encuentran en la Gerencia Regional de Aguas del Valle de México, dependiente de la SEMARNAP<sup>15</sup>, CNA<sup>16</sup>.

Los datos incluyen la ubicación, objetivo de instalación, los gastos promedio se muestran diariamente en metros cúbicos por segundo, además de un resumen de los gastos máximos, mínimos y promedios mensuales de la estación Totolica.

Para su análisis se utilizó la herramienta de SPSS<sup>17</sup> para WINDOWS ver.8.

---

<sup>15</sup> SEMARNAP. Secretaría del Medio Ambiente Recursos Naturales y Pesca.

<sup>16</sup> CNA. Comisión Nacional del Agua.

<sup>17</sup> SPSS. Paquete estadístico para la ciencia Social. Es un sistema amplio y flexible de análisis estadístico y de

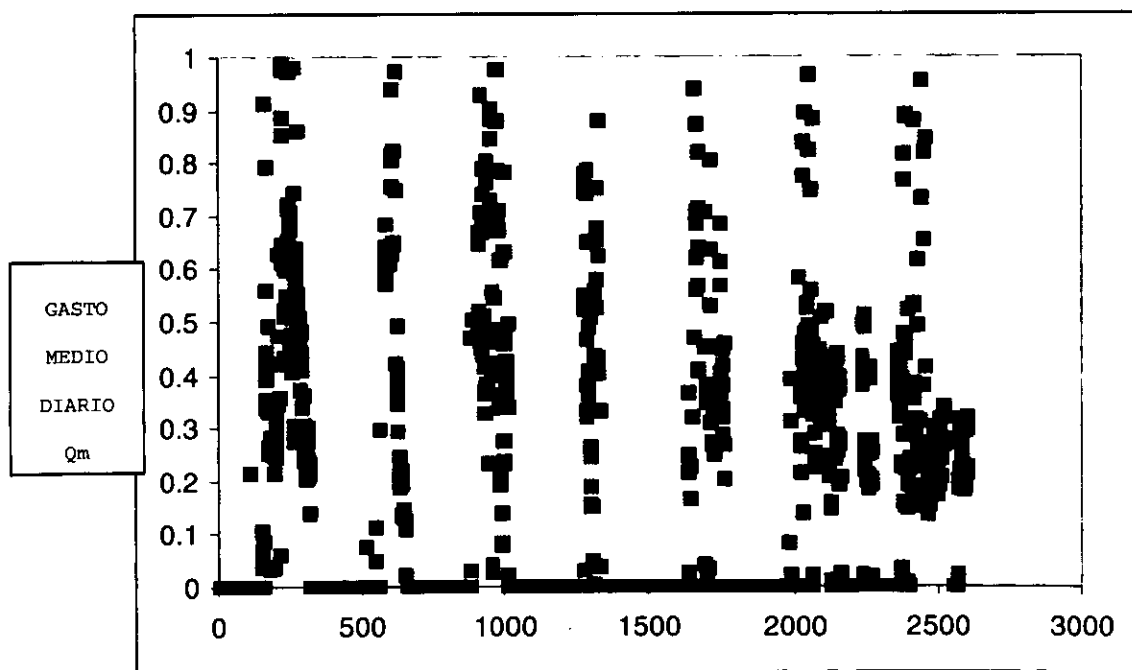
### 5.3 ANALISIS POR REGRESION LINEAL SIMPLE

Según el diccionario la palabra regresión significa "retroceso", pero estadísticamente ¿qué significa? Según Canavos las técnicas de regresión nos proporcionan medios por los cuales se pueden establecer asociaciones entre las variables de interés.

También dice que básicamente la regresión tiene dos significados:

1. Surge de la distribución conjunta de probabilidades de dos variables aleatorias.
2. Es empírico y nace de la necesidad de ajustar alguna función a un conjunto de datos.

En este caso se tratará el segundo punto. Como sólo se tiene una variable de predicción, se grafica la variable dependiente (o de respuesta) contra la variable independiente (o de predicción). En el diagrama de dispersión de los datos (ver gráfica 5.3.1) se observa un comportamiento no lineal, pero sí existe una asociación entre las variables.



SERIE ORIGINAL

TIEMPO (Días)

Gráfica 5.3.1

Al ajustar un modelo lineal simple para la variable en cuestión se tiene que la correlación es de 0.066 con tendencia positiva con una  $R^2$  de 0.004, es decir, que 4 de cada 1000 observaciones son explicadas en el modelo. Esto hace que en la hipótesis nula de que no existe el modelo sea muy débil contra la alternativa de que existe. Aplicando la prueba del estadístico Durbin-Watson<sup>18</sup> = 0.770 nos muestra que existe autocorrelación, ya que está más cerca de cero. En el análisis de la varianza (ANOVA) la  $H_0$  se rechaza y se obtiene un modelo  $Y=B_0+B_1x_1$ , donde los valores en los parámetros son  $B_0=.171$  y  $B_1=4.8(10^{-5})$  habiendo la suficiente significancia para que éstos tengan peso en el modelo.

### Reflexión

Lo anterior nos obliga a pensar 3 cosas:

1. Que para poder encontrar un modelo de regresión lineal que explique el gasto medio diario en  $m^3/seg.$  necesitamos incorporar al modelo nuevas variables explicativas que influyen en el fenómeno.
2. Generar modelos no lineales, incorporando a la misma variable explicativa con grados superiores.
3. Debido a la presencia de la autocorrelación en la variable tenemos la opción de usar otra técnica alternativa ya conocida (series de tiempo) para este tipo de casos, si no contamos con nuevas variables, que nos la minimizan o tratar de corregirla.

En la alternativa 1 no es posible incorporar nuevas variables, explorando la segunda alternativa se ajustan los siguientes modelos:

$$Y=B_0+B_1x_1^2 \quad \& \quad Y=B_0+B_1x_1+B_2x_1^2$$

<sup>18</sup> El estadístico Durbin-Watson, D, mide el grado de autocorrelación entre el residuo correspondiente a cada observación y la anterior. Si su valor es próximo a 2, los residuos estarán incorrelados, si se aproxima a 4, estarán negativamente autocorrelados, y si se aproxima a 0, estarán positivamente autocorrelados.

Modelo			A	N	O	V	A			
1	R	R <sup>2</sup>	SC	GL	MC	F		B <sub>0</sub>	B <sub>1</sub>	B <sub>2</sub>
Y=B <sub>0</sub> +B <sub>1</sub> X <sub>1</sub>	.066	.004						.171	4.8*10 <sup>-3</sup>	
	Regresión		4.697	1	4.697	12.481				
	Residual		1075.66	2858	.376					
	Total		1080.36	2859						

Modelo			A	N	O	V	A			
2	R	R <sup>2</sup>	SC	GL	MC	F		B <sub>0</sub>	B <sub>1</sub>	B <sub>2</sub>
Y= B <sub>0</sub> + B <sub>1</sub> X <sub>1</sub> <sup>2</sup>	.08793	.00773						.180999	2.12461 *10 <sup>-8</sup>	
	Regresión		173.359	2	86.6 79	231.0 89				
	Residual		1072.016 20	2858	.375 09					
	Total sin corregir		1245.369 58	2860						
	Total Corregido		1080.364 18	2859						

Modelo			A	N	O	V	A			
3	R	R <sup>2</sup>	SC	GL	MC	F		B <sub>0</sub>	B <sub>1</sub>	B <sub>2</sub>
Y=B <sub>0</sub> +B <sub>1</sub> X <sub>1</sub> +B <sub>2</sub> X <sub>1</sub> <sup>2</sup>	.11558	.01336						.299123 752	-.00021	9.01015 *10 <sup>-8</sup>
	Regresión		179.4385 8	3	59.8 12	160.3 17				
	Residual		1065.931	2857	.373 09					
	Total sin Corregir		1245.369 58	2860						
	Total Corregido		1080.364 18	2859						

Observando el cuadro anterior notamos que la R<sup>2</sup> va en aumento a medida que el modelo se complica. Se sigue cumpliendo que existe el modelo y los parámetros B<sub>0</sub>, B<sub>1</sub>, B<sub>2</sub> se vuelven más significativos en el segundo y tercer modelo.

#### Reflexión

Esto explicaría que se llegará a un buen ajuste por medio de esta técnica, a medida que incorporamos un nuevo parámetro a la variable explicativa, hasta n veces, lo que estamos haciendo es acotar los ciclos que podemos

observar desde la gráfica original, por lo tanto el encontrar un modelo por este camino resultaría muy complicado debido a que tenemos 2922 observaciones y por eso, exploremos la alternativa 3.

#### 5.4 ANÁLISIS POR SERIES DE TIEMPO

En el diccionario la palabra serie de tiempo se define como un conjunto de cosas relacionadas entre sí y que se suceden unas a otras. Matemáticamente una serie de tiempo se define como "una colección de observaciones generadas en forma secuencial a través del tiempo." <sup>19</sup>

Según T. Haan generalmente una serie de tiempo hidrológica estará compuesta por un componente estocástico superpuesto sobre un componente determinístico. Los componentes determinísticos pueden ser clasificados como periódicos, con tendencia, de brinco o una combinación de estos. La tendencia puede resultar de la naturaleza o cambios inducidos en el ambiente hidrológico. Los brincos resultan de catástrofes naturales: terremotos, grandes incendios o cambios generados por el hombre que alteran significativamente el régimen hidrológico de un área.

Los ciclos astronómicos generalmente pueden ser responsables de periodicidades en las series. Ciclos anuales son muchas veces aparentes en flujos, precipitación, evapotranspiración, etc. Ciclos semanales pueden ser presentados en el uso de agua, tales como datos sobre demanda de agua industrial, doméstica y de irrigación. Por lo anterior los datos se transformaron a un periodo de siete días. En el libro de hidrología superficial de la UNAM se dice que la tendencia es un movimiento lento de una serie en un largo periodo de tiempo; en las series hidrológicas no se toma en cuenta o bien se considera como la media de todos los valores por analizar.

Al observar la gráfica 5.3.1 de los gastos medios diarios (serie original) se nota:

Varianza: Constante	Variación aleatoria:	Si presenta
Tendencia: Constante	Fluctuación estacional:	Anual
	Modelo que sugiere:	AR(2)

<sup>19</sup> González Videgaray, María del Carmen. *Modelos de decisión con procesos estocásticos II*, UNAM-ENEP Acatlán, México, 1990; p.4.

La serie se estabilizó en varianza, como la tendencia no es muy evidente se aplicó una diferencia regular para estabilizar la media y como presentaba estacionalidad se aplicó una diferencia estacional; además de que convirtieron los datos a un periodo semanal y se obtuvieron las siguientes series:

Serie #	# Diferencia Regular	# Diferencia Estacional	Periodo
(original)			
S1	0	0	
S2	1	0	7
S3	0	1	7
S4	1	1	7

Se estimaron los órdenes de las partes regulares y estacionales de las series anteriores y se obtuvieron sus ACF's y PACF's al observarlas se ajustaron 3 series más:

Serie #	# Diferencia Regular	# Diferencia Estacional	Periodo
S5	1	1	7
S6	1	1	7
S7	1	1	7

Las series elegidas fueron las últimas 3 series (5,6,7) porque los parámetros son significativamente diferentes de cero que sugieren los modelos siguientes:

Serie 5: ARIMA(1,1,1)(2,1,1)

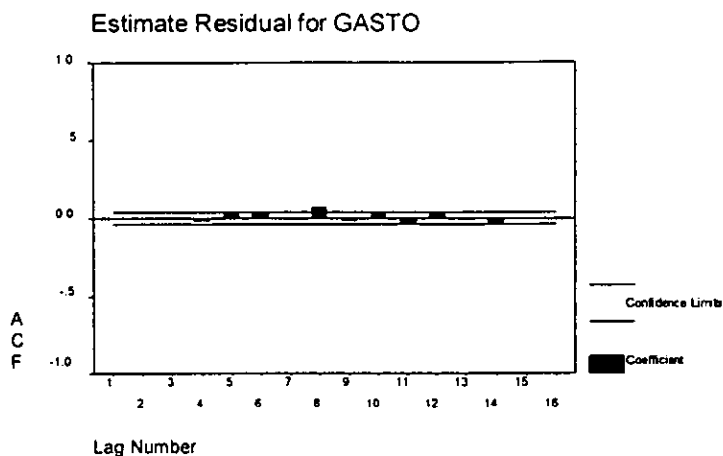
Serie 6: ARIMA(2,1,1)(1,1,1)

Serie 7: ARIMA(1,1,1)(1,1,1)

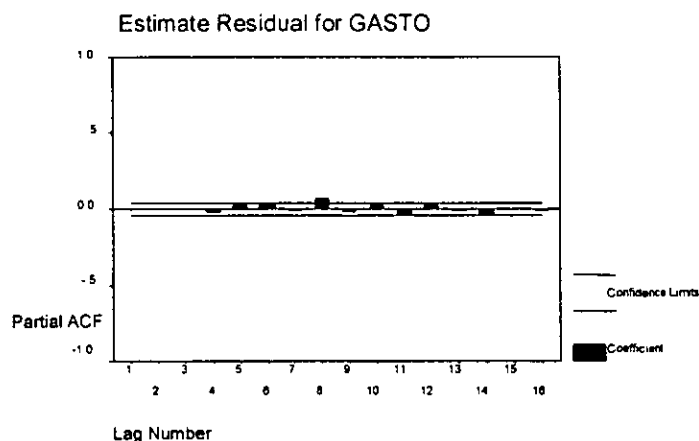
Se generaron las funciones de autocorrelación (ACF's) y las funciones de autocorrelación parcial (PACF's) de los residuos de cada modelo, aquí sólo se presentan la ACF's y la PACF's residuales de la serie S6.

#### FUNCIONES DE AUTOCORRELACIÓN DEL MODELO ELEGIDO

Gráfica 5.4.1



Función de autocorrelación residual de la serie S6.



Gráfica 5.4.2

Función de autocorrelación parcial residual de la serie S6.

Observando las gráficas 5.4.1 y 5.4.2 de las funciones de autocorrelación y autocorrelación parcial residuales del modelo elegido se nota un valor que sobresale de los límites de confianza, y no se ajusta a un comportamiento de ruido blanco, por lo que se sobreajustaron más modelos

pero se volvían más complicados a medida que se aumentaban parámetros o se efectuaban diferencias y el valor siempre estaba presente, por lo que se eligió este modelo de acuerdo al criterio AIC y SBC<sup>20</sup> y también siguiendo el principio de parsimonia<sup>21</sup>, aunque no presenta estructura de ruido blanco al tratar de mejorarlo; desde otro punto de vista, este pico nos obliga a pensar que en ése periodo pudo existir algún fenómeno<sup>22</sup> climatológico o hidrológico que influyó para que el gasto en ese año fuera alto.

TABLA DE PARÁMETROS

Serie #	S5	S6	S7
AR(1)	.35813851	.38250043	.36143547
AR(2)		.07444403	
MA(1)	.87129335	.91234974	.87351846
SAR(1)	-.08369207	-.06276323	-.07897497
SAR(2)	-.03959244		
SMA(1)	.99344989	.9903455	.99463278
Std. Error	.47023954	.4696993	.4704607
SCA*	.815.2919	779.42835	799.38775
AIC	3831.7454	3826.5085	3834.2161
SBC	3861.5243	3856.2874	3858.0392

<sup>20</sup> AIC. Akaike Information Criterion. Y SBC. Schwartz Bayesian Criterion. Generalmente el AIC es para modelos autorregresivos, mientras que SBC es un criterio más general, el mínimo AIC o SBC es el mejor; en *SPSS Trends 6.1*, p.68.

<sup>21</sup> Parsimonia: (Parsimony) este principio pretende encontrar el modelo más representativo con el menor número posible de parámetros; en González Videgaray. *Modelos de decisión con procesos estocásticos II*, p.20.

<sup>22</sup> A continuación se presenta una tabla:

## Comisión Nacional del Agua Resultados

Programa	Unidad	Periodo					
		1989	1990	1991	1992	1993	1994
<b>Condiciones Hidroclimáticas</b>							
Precipitación media anual	Milímetros	681.7	855.3	863.88	919.51	914.5	-----
Clasificación global del país		Seco	Húmedo	Húmedo	Muy húmedo	Muy húmedo	Normal
<b>Fenómenos Meteorológicos</b>							
Ciclones Océano Pacífico	Número	4	2	2	4	7	2
Ciclones Océano Atlántico	Número		1	0		2	2

Fuente: Secretaría de Agricultura y Recursos Hidráulicos. *Informe de actividades 1989-1994*; p.146. Como observamos en el cuadro anterior, el año de 1989 en general fué seco, aunque existió la presencia de 4 ciclones, uno de ellos fue el ciclón Cosme, por lo tanto podemos inferir que el gasto elevado que se muestra en las ACF's y PACF's residuales se debió a que el año 1989 fue seco y hubo necesidad de más agua para todas las actividades que dependen del agua de la presa.

\* Suma de Cuadrados Ajustados.



Se realizó el análisis de la varianza (ANOVA) de los 3 modelos elegidos.

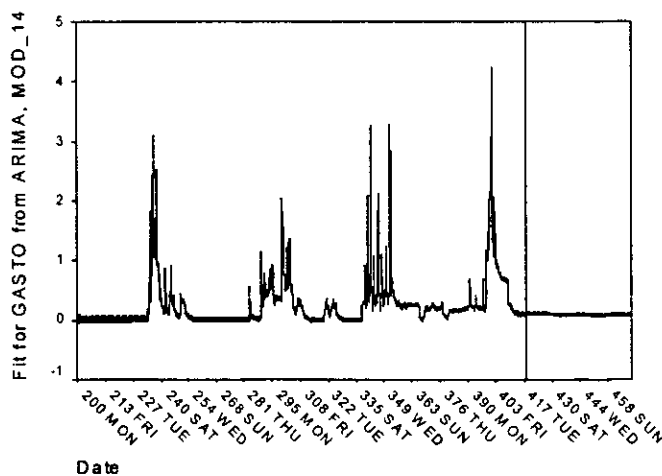
TABLA ANOVA

Modelo #	Suma de Cuadrados de residuales	Varianza residual	Grados de libertad
S5	637.69479	.22112523	2847
S6	636.52158	.22061743	2847
S7	638.69065	.22133327	2848

Modelo #	Mean		Std. Statistic	Variance Statistic	Skewness		Kurtosis	
	Statistic	Std. Error			Statistic	Std. Error	Statistic	Std. Error
S5	-1.18E-03	8.817E-03	.4708739	0.222	8.557	0.046	159.328	0.092
S6	-1.47E-03	8.810E-03	.4704785	0.221	8.631	0.046	159.648	0.092
S7	-1.16E-03	8.825E-03	.4712810	0.222	8.512	0.046	158.602	0.092

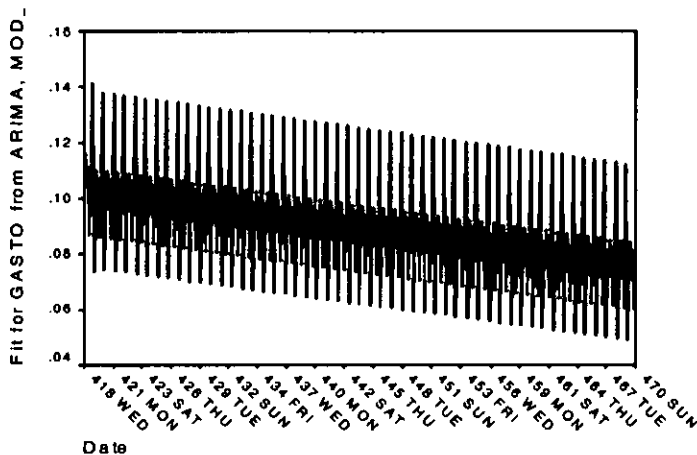
De acuerdo a la información que presentan los cuadros podemos observar que el modelo elegido es el modelo S6, porque presenta la menor varianza residual, aunque los tres modelos presentan las mismas diferencias regulares y estacionales. El modelo S6 para este caso, es el más adecuado aunque no es el óptimo.

Gráfica 5.4.3

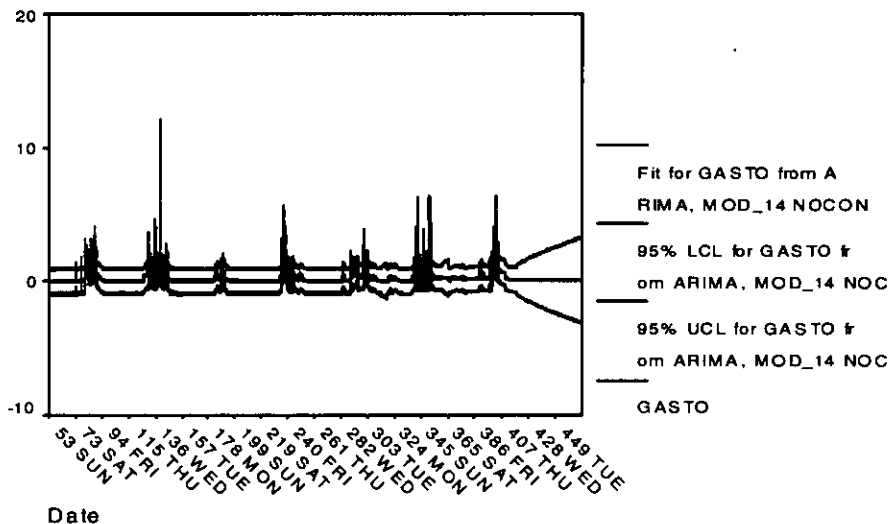


Revisando la gráfica 5.4.3 se observa una línea divisoria en la semana 417. De la semana 417 en adelante se muestra la predicción con el modelo S6.

La gráfica 5.4.4 es una ampliación del periodo de predicción de la gráfica anterior (5.4.3) que va desde la semana 418 hasta la 470 y se muestra que el gasto predicho oscila entre (0.04,0.14), lo que significa que la cantidad de agua que sale de la presa no varía mucho, aunque esta "predicción" (Fit) no es óptima y habrá que comprobarla con los datos del siguiente año (1995); que no se tienen por no estar disponibles.



Gráfica 5.4.4



Gráfica 5.4.5

La gráfica 5.4.5 muestra el periodo de validación que contiene la serie original gasto, los límites de confianza inferior y superior y la predicción generada por el modelo S6 (Fit for gasto...). También se observa que el periodo histórico que se maneja comienza en la semana 53

hasta la semana 418, que abarca el periodo 1988-1994, y de la semana 419 en adelante se muestra la predicción correspondiente al año siguiente. Se nota que la capacidad predictiva del modelo es buena porque se ajusta a los ciclos que muestra la serie original; aunque el modelo S6 es adecuado, queda por debajo del original; y la serie primitiva excede el límite superior.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

## C O N C L U S I Ó N

En el estudio de las salidas de agua de la presa Totolica, aforadas por la estación hidrométrica del mismo nombre, se comparó el ajuste de los datos mediante dos técnicas estadísticas conocidas. Para analizar los datos se usó el paquete estadístico SPSS para Windows ver.8, que es una herramienta fácil de usar, aunque un poco diferente de otras como STATGRAPHICS, SAS, etc.

Cuando se tomó la técnica de regresión lineal la bondad de la misma permitió analizar los datos en forma de prueba y error, midiendo los estadísticos interactivamente con la calibración de parámetros, pero el proceso se volvió lento y complicado a medida que se incorporaba nueva información a la variable explicativa, lo que generó modelos más complejos, influyendo a la misma variable endógena hasta llegar a los modelos no lineales en grados superiores provocando acotamiento en su comportamiento del fenómeno jugando el papel decisivo en este ajuste determinístico. Aunque se realizó con tecnología esta mecánica, lo que se hace es encontrar un modelo polinomial que se vaya ajustando de acuerdo a  $n-1$  observaciones que se tengan.

La técnica de series de tiempo se adaptó mejor al conjunto de datos, y se obtiene el modelo con mayor rapidez, aunque modelar datos diarios es un proceso más complicado y es necesaria más experiencia el trabajar con series hidrológicas, por ser fenómenos aleatorios.

Se encontró un modelo adecuado  $ARIMA(2,1,1)(1,1,1)$ , pues se ajusta a los ciclos que tiene la serie original, aunque las gráficas de los residuales no mostraron un comportamiento de ruido blanco, porque se observa un valor que excede los intervalos de confianza.

Se ajustaron varios modelos más para tratar de eliminar ése pico, pero en todos los modelos probados se seguía presentando el mismo valor, por lo

que esta situación llevó a pensar que hubo algún fenómeno climático o hidrológico que influyera de alguna forma. Buscando información se encontró que el año de 1989 fue un año seco, por lo que esto pudiera ser una causa del elevado gasto. Por lo tanto se escogieron las series S5:ARIMA(1,1,1)(2,1,1), S6:ARIMA(2,1,1)(1,1,1), S7:ARIMA(1,1,1)(1,1,1) donde el valor es menor, y S6 es el modelo elegido.

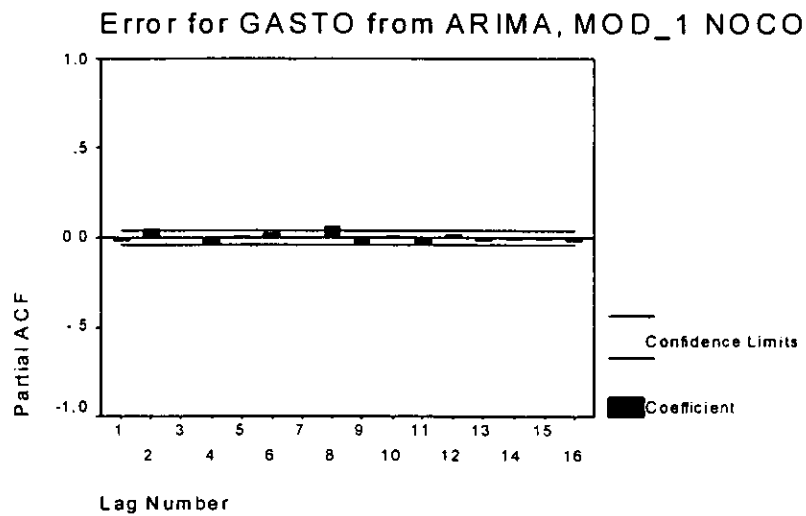
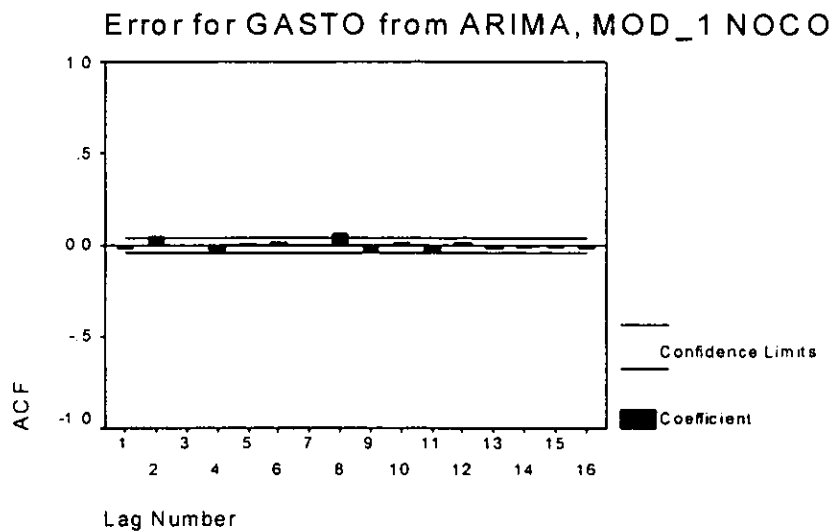
Reflexionando, se puede decir que no se encontró un modelo aceptable porque quizá no se manejó correctamente la información, es decir, se trabajó con gastos diarios transformados a un periodo semanal, porque datos sobre demanda de agua pueden ser presentados en ciclos semanales y pienso que es más fácil trabajar con datos mensuales, reduciéndose el error, cuando sacamos la media de estos gastos; es necesaria más experiencia en el manejo e interpretación de los datos además de que generalmente la inexactitud en las predicciones se debe a:

- errores en la medición o captura de los datos,
- al modelo utilizado, o
- los datos que varían rápidamente por ser fenómenos imprevistos.

Comparando las dos metodologías empleadas, se concluye que son buenas herramientas de análisis y aplicación, las dos nos proporcionan un modelo con el cual se puede predecir el fenómeno en estudio; por igual nos permiten estimar la variable dependiente de la variable independiente; también nos permiten describir el comportamiento del fenómeno. Pero para este caso, se ajustó mejor al conjunto de datos la técnica estadística de series de tiempo.

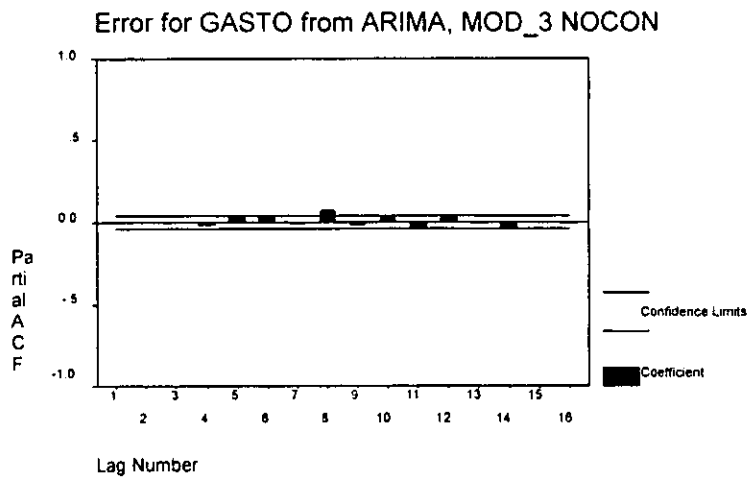
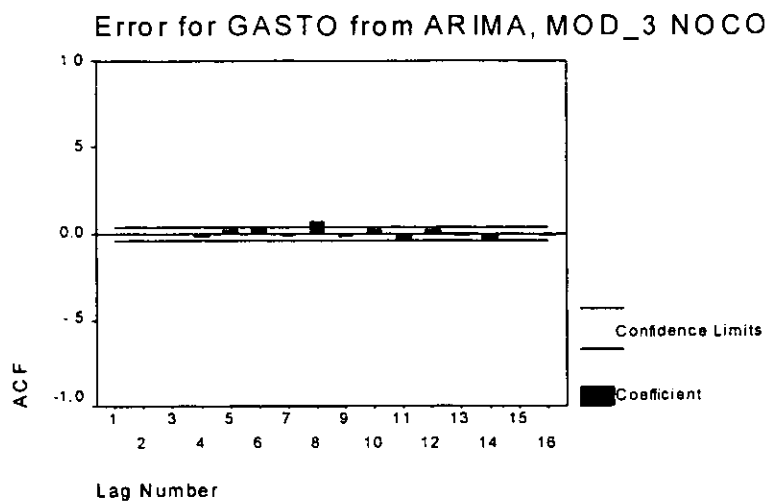
Por último se cumplió el objetivo al comparar el ajuste de los datos mediante las dos técnicas, éstas nos explicaron el comportamiento del gasto a través del tiempo y se acertó en la consideración de que el ajuste fue mejor con la serie cronológica. Además de que esta investigación sirvió para tener una visión más amplia de las aplicaciones de las técnicas estadísticas y la aplicación de la licenciatura en otras áreas del conocimiento.

## ANEXO 1

G R A F I C A S  
ACF Y PACF

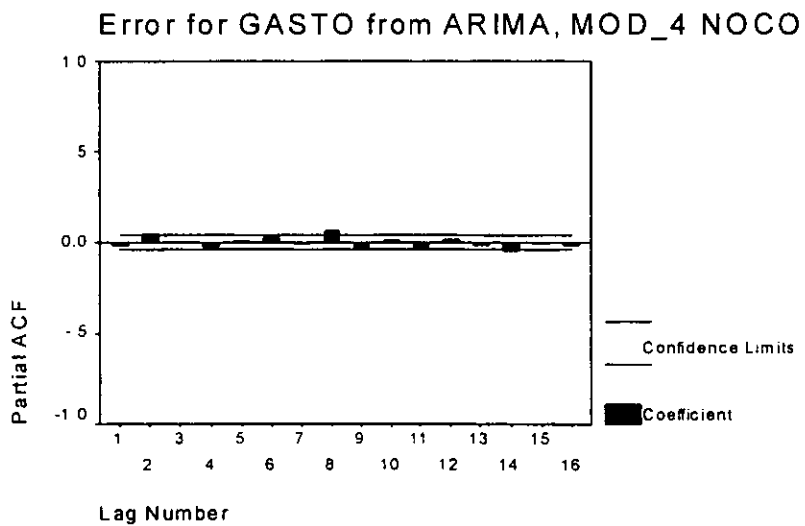
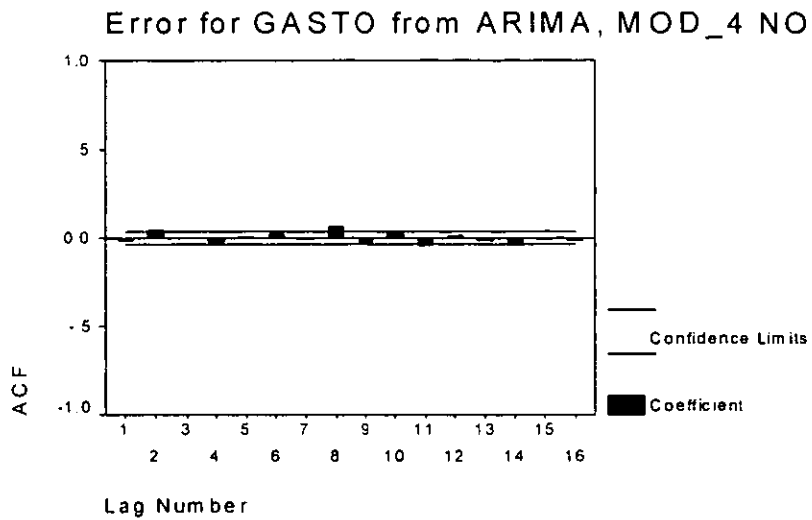
Gráfica 1. Función de autocorrelación de la serie S5.

Gráfica 2. Función de autocorrelación parcial de la serie S5.



Gráfica 3. Función de autocorrelación de la serie S6.

Gráfica 4. Función de autocorrelación parcial de la serie S6.



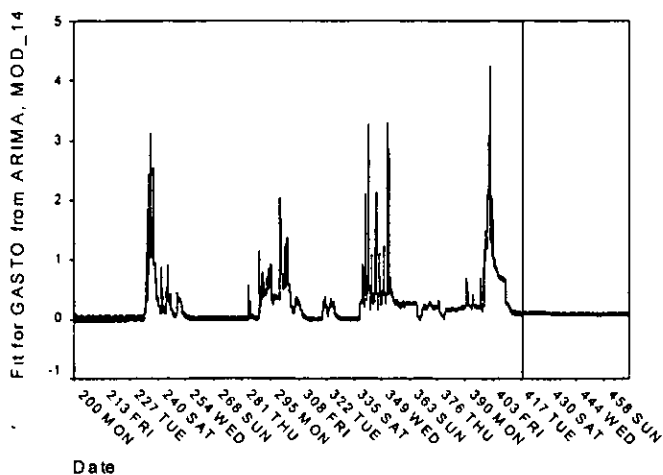
Gráfica 5. Función de autocorrelación de la serie S7.

Gráfica 6. Función de autocorrelación parcial de la serie S7.

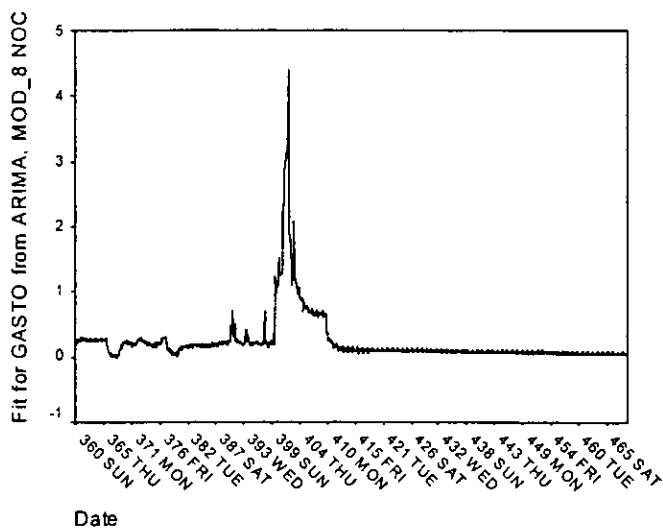
Observando las gráficas de las funciones de autocorrelación y autocorrelación parcial residuales de los 3 modelos elegidos se nota un valor que excede de los límites de confianza, se ajustaron más modelos pero se volvían más complicados a medida que se aumentaban parámetros o se efectuaban diferencias y el valor siempre estaba presente.



Gráfica (7) de predicción para un año de la serie S6. De la semana número 418 a 458 se muestra la "predicción" para el año siguiente.

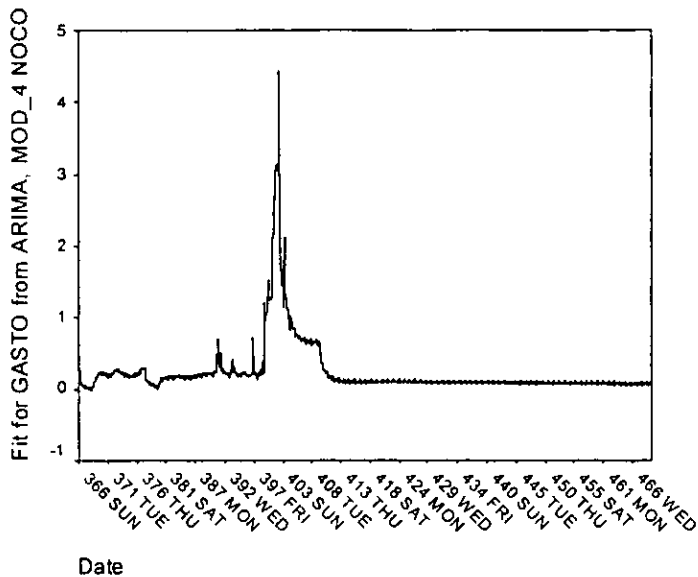


Gráfica 7



Gráfica 8

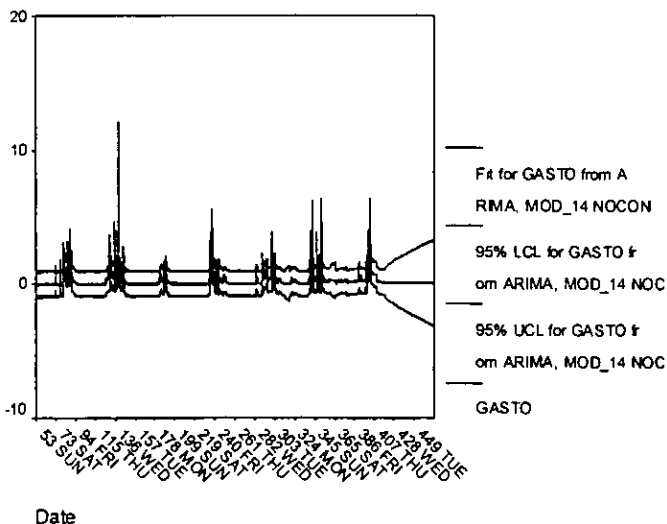
Gráfica (8) de predicción para la serie S5. De la semana 418 a la 466 se muestra la "predicción" con este modelo para el año de 1995.



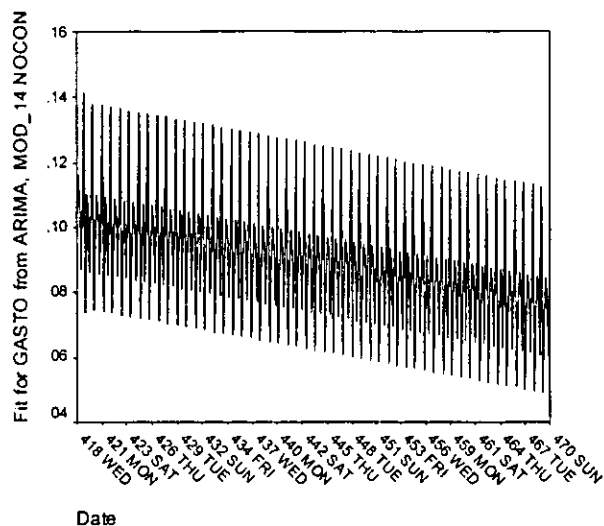
Gráfica 9

Esta es la gráfica (9) que muestra la predicción con la serie S7, la predicción corre desde la semana 418 a la 466, (que forman un año).

Revisando las gráficas del periodo de predicción de los tres modelos se puede observar que el gasto siempre está sobre el cero, lo que significa que la cantidad de gasto que sale de la presa es casi siempre la misma, aunque esta "predicción" no es óptima y habrá que comprobarla con los datos del año siguiente (1995); que no se tienen por no estar disponibles.



Gráfica 10



Gráfica 11

En la gráfica 10 se muestran los límites inferior y superior, la serie original *gasto* y la predicción generada por la serie *S6(Fit...)*, en la que se observa un ajuste apropiado con el modelo elegido, aunque el *gasto* sobrepasa del límite superior.

La gráfica 11 presenta una ampliación del periodo semanal (418-470) de las predicciones (Fit) durante el año siguiente (1995), como se observa las predicciones oscilan entre (0.04,0.14).

En la siguiente página se muestran los nombres de las corrientes, presas, estaciones hidrométricas, corrientes entubadas y canales del mapa 3 del capítulo 4.

**NOMBRES DE ESTACIONES HIDROMETRICAS, PRESAS, ETC. DEL MAPA 3**

**CORRIENTES**

1.A. S.Gregorio  
 2.R. S. Lucas  
 3.A. Santiago  
 4.R. S.Buenaventura  
 5.R. s. Juan de Dios  
 6.R. Eslava  
 7.R. Magdalena  
 8.Barranca Anzaldo  
 9.R. Texcalatlaco  
 10. B. De Guadalupe  
 11.B. del Muerto  
 12.R. Mixcoac  
 13.R. Becerra  
 14.R. Tacubaya  
 15.Barranca de dolores  
 16.Barranca Barrilado  
 17.R. Tecamachalco  
 18.R. s. Joaquín  
 19.Barranca del Tornillo  
 20.A. Agua Blanca  
 21.R. Huixquilucan  
 22.A. Agua de los leones  
 23.R. Chinaco  
 24.R. del Borracho  
 25.R. Sordo  
 26.R. los Cuartos  
 27. R. Totolica  
 28.R. Hondo  
 29.R. Chico Remedios  
 30.A. Colorado  
 31.A. S. Mateo  
 32.A. Castro  
 33.A. Madín  
 34.A. Sifón  
 35.A. Chiluca  
 36.A. Alameda  
 37.A. de Córdoba  
 38.R. Tlalnepantla  
 39.R. S. Javier  
 40.R. Cuautepec  
 41.A. Guadalupe  
 42.A. de la Colmena  
 43.A. Bata  
 44.A. Monte Alto  
 45.A. Chiquito  
 46.R. San Pedro  
 47.A. San Miguel  
 48.Barranca Lanzarote  
 49.Barranca Alcaparrosa  
 50.R. Tepetzotlán  
 51.R. de Avs. de Pachuca  
 52.R. Papalotla  
 53.A. Azoyatla  
 54.A. La Palama  
 55.R. S. Juan Teotihuacán  
 56.Barranca de Maquixco  
 57.A. San Martín  
 58.Barranca Atlamajac  
 59.Barranca los Muertos  
 60.B. Tlachichinámitl  
 61.Barranca Metepec  
 62.R. Papalotla  
 63.Barranca de las musas  
 64.A. Piedras Negras  
 65.Barranca Xanco  
 66.A. Xalistaca  
 67.R. Zalapango

68.A. Sta. Inés  
 69.A. Comunidad  
 70.A. Purificación  
 71.R. Coxacoaco  
 72.A. Magdalena  
 73.R. Texcoco  
 74.R. Chapingo  
 75.R. S. Bernardino  
 76.R. Sta. Mónica  
 77.Cañada del Agua  
 78.A. Potreritos  
 79.A. Tejocote  
 80.A. Caño Viejo  
 81.R. Coatepec  
 82.A. Puente Quebrado  
 83.A. Metepango  
 84.A. Las Mojadas  
 85.A. Costitlán  
 86.R. S. Francisco  
 87.A. Tzintzintla  
 88.Barranca Sta. Cruz  
 89.Cañada S. Francisco  
 90.Barranca Cantera  
 91.Barranca Tecolotte  
 92.A. Potrero  
 93.A. Santo Domingo  
 94.Barranca S. Domingo  
 95.Barranca Popotla  
 96.R. de la Compañía  
 97.Barranca Huejolaca  
 98.Barranca paso del ja  
 99.Cañada del Negro  
 100.A. Tlalmanalco  
 101.Cañada Alcalicán  
 102.A. San José  
 103.A. Paraje  
 104.A. de los Diamantes  
 105.A. San Antonio  
 106.A. Panchaya  
 107.R. San Luis  
 108.R. Coacalco  
 109.Cañada de Oluca  
 110.R. Ameca  
 111.A. Zacualtitlac  
 112.A. Tenechcolux  
 113.A. Zacatlaco  
 114.R. Milpa Alta  
 115.A. Vaquería  
 116.A. El Columpio  
 117.A. Calpulalpan  
 118.A. El Tizar  
 119.R. San José  
 120.Barranca del Muerto  
 121.R. San Miguel  
 122.R. Coatlico  
 123.R. El canal

**ESTACIONES****HIDROMÉTRICAS**

(1) Desviación alta al pedregal. R. Magdalena  
 (2) Totolica.Río Totolica  
 (3) Molinito.R.Hondo  
 (4) El Salitre.R.Tepetlaxco

(5) Molino Blanco.R. Los Remedios  
 (6) Km0+393. Emisor Poniente  
 (7) Calacoaya.R. Tlalnepantla  
 (8) Las Arboledas. R. San Javier  
 (9) San Martín Obispo, Emisor poniente  
 (10) San Lorenzo.R. Cuautitlán  
 (11) San Lorenzo, Emisor Poninete  
 (12) Puente Colgante. R. Cuautitlán  
 (13) Huhuetoca. R. Cuautitlán  
 (14) El Tajo de Tequixquiac  
 (15) Tunnel nuevo de tequixquiac  
 (16) El manantial. R. avenidas de Pachuca  
 (17) Km27+250. Gran Canal  
 (18) Km2+120 Canal de Desfogue  
 (19) Km9+500 Desviación combinada  
 (20) Km5+250granCanal  
 (21) Puente el Arenal  
 (22) Vertedor Milpa Alta  
 (23) S. Luis R. Amecameca  
 (24) S. Lucas. R. la Compañía  
 (25) S. Marcos.R. S.Fco.  
 (26) El Tejocote. Sta.Mónica  
 (27) S.Mateo .R. S.Bernardino  
 (28) Chapingo, R. Chapingo  
 (29) Texcoco, R.Texcoco  
 (30) S.Andrés,R. Coxacoaco  
 (31) Atenco, R.Xalapango  
 (32) La Grande.R. Papalotla  
 (33) Tepexpan.R. S.J.Teotihuacán  
 (34) Irolo,Canal Salidas Laguna de Apan  
 (35) S.Jerónimo.C. Desagüe  
 (36) Potrerito.R.Barranca del Monte  
 (37) S. Marcos, R. Tizar.

**PRESAS**

I. Desviación Alta  
 II. Anzaldo  
 III. S.Jerónimo

IV. Texcalatlaco  
 V. Atlamaya  
 VI. Tarango  
 VII. Mixcoac  
 VIII. Becerra  
 IX. Tacubaya  
 X. Dolores  
 XII. Tecamachalco  
 XIII.S.Joaquín  
 XIV. El Tornillo  
 XV. El Sordo  
 XVI. Los Cuartos  
 XVII. Las Julianas  
 XVIII. los Arcos  
 XIX. El Colorado  
 XX. La Colorada  
 XXI. Madín  
 XXII. Vaso el Cristo  
 XXIII. Vaso Carretas  
 XXIV.S. Mateo  
 XXV.S. Joaquín  
 XXVI. Las Ruinas  
 XXVII. Rancho viejo  
 XXVIII. La Colmena  
 XXIX. Agustin de Iturbide  
 XXX. Guadalupe  
 XXXI. Angulo  
 XXXII. El Tesoro  
 XXXIII. Rancho Fuentes  
 XXXIV. El Muerto  
 XXXV. La Piedad  
 XXXVI.S. Miguel  
 XXXVII. La Concepción  
 XXXVIII. S.Guillermo  
 XXXIX. Cuevecillas  
 XL. El Girón.

**CORRIENTES****ENTUBADAS Y****CANALES**

A.Tunnel del Acueducto de Lerma.  
 B.Desviación Combinada.  
 C.Río consulado  
 D.R. de la Piedad.  
 E.R.Churubusco.  
 F.Interceptor Poniente.  
 G.interceptor Central  
 H.interceptor Oriente  
 I.Gran Canal del Desagüe  
 J.Tajo de Nochistongo.  
 K.Nuevo tunnel de Tequisquiac  
 L. Antiguo tunnel de Tequisquiac.  
 M.Canal de Castro.  
 N.Canal de Sales  
 O.Bordo Poniente.  
 P. Bordo de Xochiaca.  
 Q. Canal Nacional.  
 R. Canal de Garay.  
 S. Canal de Chalco.  
 T. Canal de Cuernavaca.  
 U. Canal de Caltongo.  
 V. Canal General.



## ANEXO 2

### CONCEPTOS DE HIDRODINAMICA

**HIDRODINAMICA:** Estudia el comportamiento de los líquidos en movimiento. Estudia el comportamiento del agua en movimiento en las obras o maquinas de ingeniería.

**CAUDAL:** Gasto o caudal es un volumen dividido por un tiempo. Es el volumen que por unidad de tiempo pasa por un lugar determinado (por una sección determinada) suele medirse en m<sup>3</sup>/seg o en litros por segundos. Caudales muy pequeños se miden a veces en litros por minutos o en litros por hora; m<sup>3</sup>/seg es lo mismo que decir: metros cúbicos de agua que pasan en un segundo por una sección normal a la corriente.

Debemos distinguir los conceptos de caudal medio y caudal instantáneo.

**CAUDAL MEDIO:** Si medimos todo volumen de agua escurrido en cierto tiempo por una sección determinada y lo dividimos por todo el tiempo que ha tardado en escurrir obtendremos el caudal medio:

$$\text{caudal medio} = Q_m = \text{volumen total} / \text{tiempo total} = v/t$$

Inversamente, suponiendo un caudal constante  $Q$ , el volumen total escurrido en un tiempo  $t$  determinado será igual al producto del caudal por el tiempo  $t$ :

$$v = Q \times t$$

**CAUDAL INSTANTANEO:** =  $Q$  = volumen pequeño escurrido/tiempo muy pequeño

Para calcular el volumen total de agua escurrido en cierto tiempo por una sección de una corriente cuyo caudal sea variable se divide el tiempo total en varios intervalos (iguales o no según sea más cómodo o mas adecuado) de corta duración. Durante cada pequeño intervalo habrá estado escurriendo un caudal que puede suponerse constante en ese intervalo (o caudal medio de ese intervalo).

El volumen total será:

$$V = Q_1t_1 + Q_2t_2 + Q_3t_3 + \dots + Q_nt_n = S(Q t)$$

en que  $Q_1, Q_2, \dots, Q_n$  son dichos caudales representativos o medios durante cada pequeño lapso de tiempo respectivo  $t_1, t_2, \dots, t_n$ .

**LIMNIGRAFO:** consta de una aguja con lápiz o tinta que incide sobre un rollo de papel especialmente rayado, que proviene de un cilindro y se enrolla de nuevo en otro. El papel se va desplazando por medio de un aparato de relojería bien sincronizado. De este modo queda constancia sobre el papel, de todos los niveles habidos a través del tiempo. Como estos niveles son lo de una sección transversal conocida, sirven para calcular las lecturas de agua o calados.

Cuando sea el caso, con auxilio de la curva de descarga de esa sección conoceremos el caudal  $Q$  que estuvo pasando por el río o canal en cada momento. Las escalas de papel suelen ser de los 3mm por hora y para los niveles, desde 1:1: a 1:50 según los casos. Una escala usual, en canales abiertos es la de 1:5. Cuanto mayor es la escala, mayor apreciación puede hacerse, pero los gráficos pueden salirse del papel. Existen, sin embargo, aparatos en los que cuando la aguja llega a uno de lo bordes se desplaza bruscamente al otro borde, para aprovechar de nuevo el papel de este modo pueden usarse escalas grandes sin temor a salirse de la hoja de papel.

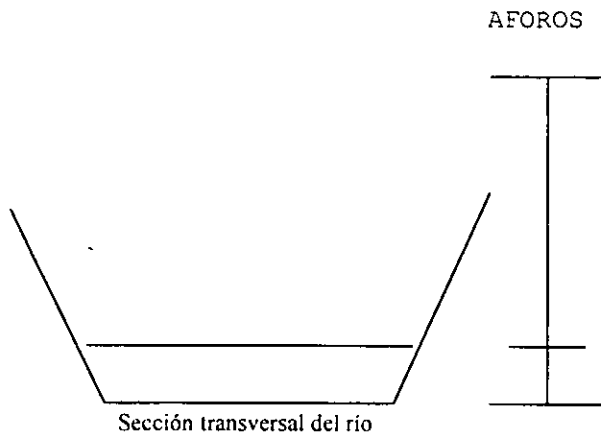
Lo normal es que los tiempos están horizontales (a lo largo del rollo y las alturas de agua en verticales (a través del papel). Los datos de los limnigrafos se usan para llevar estadísticas de los caudales de corrientes naturales y canales.

**VOLUMEN ESCURRIDO EN UN DETERMINADO TIEMPO.** A veces es necesario conocer el volumen total escurrido por un canal o río en un tiempo determinado, frecuentemente un año. Para ello hay que conocer el caudal medio diario  $Q_m$ , día a día, dato que se obtiene de la práctica de aforos, ya sea con limnigrafo o por lecturas de escala (sistema menos preciso). El problema es simple, pues:

Volumen total diario en  $m^3 = Q_m \times 24 \times 60 \times 60 = 86400Q_m$  viniendo  $Q_m$  en  $m^3/\text{seg}$ ,

Si disponemos de datos de limnigrafos, debemos calcularlos (a través de las cotas de nivel del agua) la curva de caudales instantáneo de la que por integración gráfica o con sencillos programas de microordenador podremos calcular los volúmenes diarios.

Basta sumar los volúmenes diarios, para obtener el volumen total en el tiempo que sea.



$$60 * 60 * 24 = 86400 \text{ seg.} / 1000000 = 0.086400$$

Esto es 1min. = 60 seg.

1hr. = 60 min.

1 día = 24 horas

Ejemplo:  $Q = \text{m}^3/\text{s} = 1000 \text{ L} / \text{s}.$

$$= 1 \text{ m}^3/\text{s}$$

$$9 \text{ m}^3/\text{s} * 0.0864 = 0.777\text{Mm}^3 \text{ Volumen}$$

**MEDICION DE CAUDALES (AFOROS):** Se llama **aforar** una corriente, calcular el gasto  $Q$  que escurre. Hay varios procedimientos algunos son:

Aforos por el método volumétrico, con flotadores, mediante las fórmulas empíricas de canales y tuberías, por orificios y vertederos, por disolución comparada (aforos químicos), por ondas coloreadas, por pantalla móvil, con tubos Venturi, por el procedimiento de Gibson, por el procedimiento de Allen, etc.



**MEDICION DE VELOCIDADES CON MOLINETE:** existen dos clases de molinetes: de cazoletas y de hélice. El molinete de cazoletas lleva un anillo de cazoletas cónicas que giran alrededor de un eje vertical montado en el cuerpo que se sumerge en la corriente hasta el punto donde se requiere medir la velocidad.

El molinete de hélice tiene como su nombre lo indica una hélice de eje horizontal, la que es movida por el paso del agua, este es mas preciso que el de cazoletas, en corrientes muy turbulentas es más recomendable. Tanto uno como otro se estabilizan girando a velocidad constante. Cada aparato dispone de una tabla curva o ecuación que relaciona el numero de revoluciones por segundo con la velocidad  $v$  del agua.

Frecuentemente la ley es lineal, para velocidades ordinarias:  $v = K \cdot n$  en donde:

$v$  = velocidad del agua en m/seg.

$k$  = una constante del aparato,

$n$  = número de vueltas por segundo,

Para calcular  $n$  sin equívocos, los aparatos disponen de unos contactos eléctricos alimentados por una pila, que hace sonar un timbre cada 10 vueltas. El observador totaliza el numero de vueltas  $N$  y el tiempo  $t$  segundos en que se han realizado calculando  $n = N/t$

No hay que contar desde el principio: conviene dejar pasar unos cuantos segundos para que el molinete coja su velocidad de régimen.

Se opera con un cronómetro cuenta-segundo y no con un reloj ordinario. Las medidas se repiten por lo menos dos veces, mas si se observa algún resultado extraño o poco concordante con el anterior. Dejando el molinete en diferentes puntos de una vertical se van obteniendo las velocidades a diferentes profundidades; el vástago del molinete sirve de sonda e indica las metas de inmersión del aparato. Hay que cuidar que la sonda este vertical (molinete horizontal). Para usar la ecuación anterior pueden seguirse dos procedimientos:

Contar el número  $N$  de vueltas en un tiempo exacto, (100 seg. por ejemplo) o ver el tiempo en que se realizan 100 (ó 200) vueltas.

**ANEXO 3****G L O S A R I O**

- Afluente:** Río o arroyo que desemboca a otro mayor.
- Aforar:** Medir la cantidad de agua que lleva una corriente en una unidad de tiempo.
- Aforo:** Medida de la cantidad de agua.
- Aguas abajo:** Hacia la parte inferior de una corriente.
- Aguas arriba:** Hacia la parte superior de una corriente.
- Atarjea:** Cañería, alcantarilla, tajea, alcantarilla para el paso del agua debajo de los caminos.
- Avenida:** ver crecida.
- Barlovento:** Lado de donde procede el viento.
- Cauce:** lecho de un río, canal o acequia. Conducto abierto por donde fluye la corriente.
- Cauce inferior:** Parte del río donde disminuye el declive y también la erosión.
- Cauce superior:** Parte del río donde la acción erosiva va en aumento.
- Caudal:** Cantidad de agua que mana o corre; volumen de agua de los manantiales o de los ríos. Es un volumen dividido por un tiempo. Es el volumen que por unidad de tiempo pasa por un lugar determinado, suele medirse en m<sup>3</sup>/seg. Ver gasto.
- Corriente o Escurrimiento:** es el estado de movimiento del agua.
- Crecida o Avenida:** Se llama crecida o avenida al hecho de que eventualmente, un río presente, durante un tiempo relativamente corto un caudal muy grande. Las crecidas son un problema importante porque pueden destruir las obras de ingeniería, inundar terrenos, poblaciones y provocar víctimas.
- Cuenca:** En general se da este nombre a cualquier depresión de la litosfera, este cubierta o no por agua. La cuenca de un río es toda el área drenada del río y sus tributarios. Territorio cuyas aguas escurren a un mismo río, lago o mar.

- Depresión:** Area de la superficie terrestre baja o hundida.
- Divisoria:** Línea que puede considerarse idealmente en una región, y desde la cual las aguas corren en direcciones opuestas.
- Efluente:** Este término es usado generalmente para la descarga de aguas residuales a ríos, lagos o cuerpos de agua en general.
- Embalse:** Retención artificial de las aguas de un río para utilizarlas en la producción de energía o en el riego de los campos.
- Esclusa:** Recinto en un canal de navegación con puertas movibles de entrada y salida que se pueden cerrar y abrir según se quiera contener las aguas o dejarlas correr.
- Escurrimiento:** Flujo superficial del agua proveniente de un área o volumen total del flujo durante un tiempo especificado.
- Estación automática:** Estación en la que los instrumentos efectúan y transmiten o registran automáticamente las observaciones, realizando, en caso necesario la conversión a la clave correspondiente o bien realizándose esa conversión en una estación transcriptoras.
- Estación hidrológica:** Estación en la cual se obtienen datos del agua en los ríos, lagos y embalses, de uno o varios de los elementos siguientes: niveles, flujo de corrientes, transporte y depósito de sedimentos, temperatura del agua y otras propiedades químicas del agua.
- Estación sinóptica:** Estación desde la cual se efectúan observaciones sinópticas.
- Estiaje:** Epoca del año en que los cauces tienen escurrimiento mínimo o se secan por completo. Disminución del caudal de los ríos debido a la sequía.
- Estuario:** Desembocadura profunda de un río, donde se mezclan las aguas del mar y del río.
- Evapotranspiración:** pérdida de agua del suelo por evaporación y transpiración del proceso de crecimiento de las plantas en el suelo.
- Gasto:** Volumen de agua que pasa por la sección transversal de un conducto o río en la unidad de tiempo; se mide en m<sup>3</sup>/seg, por minutos, por horas, por día, también suele expresarse en litros por segundo, por minutos, etc.
- Hidrodinámica:** Estudia el comportamiento del agua en movimiento. Estudia el comportamiento del agua en movimiento en las obras de ingeniería.
- Hidrograma:** Gráfico que muestra la variación con respecto al tiempo del nivel de agua, del caudal o de la velocidad u otras características de las corrientes de agua.

- Hidrometría:** Parte de la hidrología que trata de la medición del agua que escurre por las corrientes o que está contenida en depósitos naturales o artificiales.
- Isoyetas:** Líneas que en los mapas unen puntos de igual precipitación pluvial.
- Lecho:** Cauce, fondo.
- Limnógrafo:** Consta de una aguja con lápiz o tinta, que incide sobre un rollo de papel especialmente rayado que proviene de un cilindro y se enrolla de nuevo en otro.
- Observación hidrológica:** Medición directa o evaluación de uno o varios elementos hidrológicos, tales como el nivel del agua, el caudal, la temperatura del agua, etc.
- Precipitación pluvial:** Caída del agua. Lluvia.
- Predicción climatológica:** Predicción basada en el clima de una región.
- Reducto:** Fortificación, refugio, posición.
- Régimen:** Modo de hacerse el escurrimiento de las aguas de un río.
- Sequía:** Tiempo seco de larga duración.
- Sotavento:** Porción opuesta al lugar de donde viene el viento.
- Tajo:** Corte profundo en el terreno.

**REFERENCIA BIBLIOGRAFICA**

Bribiesca Castrejón, José Luis. "Hidrología histórica del Valle de México". Rev. *Ingeniería Hidráulica en México*. Volumen XIV, núm. 3 julio-septiembre 1960, México, pp.43-59.

Canavos, George C. *Probabilidad y Estadística Aplicaciones y Métodos* (Traductor Urbina Medal, Edmundo G.) Edit. McGraw-Hill/Interamericana de México, S.A. de C.V., México, 1988, 65lp.

Comisión Nacional del Agua (CNA). *Instructivo para aforo de corrientes*. 7ª ed., México, Subdirección General de Administración del Agua, Gerencia de Aguas Superficiales e Ingeniería de Ríos, 1993, 230p.

Escriba Bonafé, Domingo. *Hidráulica para ingenieros*. 1ª ed., Librería Editorial Bellisco, Madrid, 1988, 684p.

Ezcurra Exequiel. *De las Chinampas a la Megalópolis: el medio ambiente en la cuenca de México*. SEP, FCE, CONACYT, (La Ciencia desde México/91), 1995, 119p.

Ferrán Aranz, Magdalena. *SPSS para Windows programación y análisis estadístico*. McGraw-Hill/Interamericana de España, S.A., Aravaca, Madrid, España, 1996, 580p.

González Videgaray, María del Carmen. *Modelos de decisión con procesos estocásticos II: metodología Box-Jenkis*. UNAM-ENEP Acatlán, México, 1990, 114p.

Helweg J., Otto. *Recursos Hidráulicos: Planeación y Administración*. Versión en español Carlos H. García Fener, Edit. Limusa, Grupo Noriega Editores, México, 1976, 405p.

Instituto de Geografía, UNAM. *Atlas Nacional de México*. Volúmen II, 1990.

INEGI. *Anuario Estadístico del Estado de México*. México, edición 1997.

\_\_\_\_\_. *Estadísticas del Medio Ambiente*. 1ª ed. marzo de 1995, México, 1994, 447p.

Jáuregui Ostos, Ernesto. "Los Huracanes prefieren a México". *Información Científica y Tecnológica*. Revista mensual, México, D.F., Agosto de 1989, vol.11; núm. 155, 70p.

M. Shaw, Elizabeth. *Hidrology in Practice*. 3ª ed., Editorial Chapman & Hall, U.S.A., 569p.

Neter John, Kutner H. Michael, Nachtsheim J. Christopher, Wasserman William. *Applied Linear Statistical Model*. 4ª ed. U.S.A., IRWIN 1983, y 1989, Times Mirror Higher Education Group Inc. 1996., 1408p.

Organización Meteorológica Mundial. *Guía de Prácticas Hidrológicas*. Volúmen I y II, OMM, Ginebra, Suiza, 1990.

Sáenz de la Calzada, Carlos. *Geografía General*. 1ª ed. México, D.F. Editorial Esfinge, México, 1977, 288p.

Sánchez Molina, Antonio. *Síntesis Geográfica de México*. 10ª ed., México., Editorial Trillas, 1962, 265p.

Secretaría de Agricultura y Recursos Hidráulicos. *Informe de Actividades 1989-1994*. México, D.F. noviembre 1994, 185p.

Secretaría del Medio Ambiente Recursos Naturales y Pesca, Gerencia Regional de Agua del Valle de México. *Boletín Hidrológico datos del Valle de México*. Varios tomos (1987-1994).

SPSS Inc. *Spss Trends 6.1*. U.S.A., 1994, 355p.

T. Haan, Charles. *Statistical Methods in Hidrology*. The Iowa State University Press, U.S.A., 1977, 500p.

UNAM. *Manual para el aforo de Corrientes*. Facultad de Ingeniería. División de Estudios Superiores, SRH. 1970, 396p.

\_\_\_\_. *Análisis Estadístico y Probabilístico de Datos Hidrológicos. Hidrología superficial*. Facultad de Ingeniería, México, D.F. 50p.

#### OBRAS CONSULTADAS

Barnett, Vic & Feridun Turkman, K. ed. *Statistics for Environment*. John Wiley & sons, Lisboa, Portugal, 1993, 427p.

Lloy, E. H., O'Donnell, T., Wilkiwson, J. *The Mathematics of Hidrology and Water resources*. Academic Press, N.Y., New York, 1979, 138p.