



RENIECYT - LATINDEX - Research Gate - DULCINEA - CLASE - Sudoc - HISPANA - SHERPA UNIVERSIA - E-Revistas - Google Scholar
DOI - REBID - Mendeley - DIALNET - ROAD - ORCID

Title: Minado de series de tiempo utilizando la metodología ARIMA

Author: Gil Santana-Esparza

Editorial label ECORFAN: 607-8324
BCIERMIMI Control Number: 2016-01
BCIERMIMI Classification(2016): 191016-0101

Pages: 17
Mail: *gil.santana@itspanuco.edu.mx*
RNA: 03-2010-032610115700-14

ECORFAN-México, S.C.

244 – 2 Itzopan Street
La Florida, Ecatepec Municipality
Mexico State, 55120 Zipcode
Phone: +52 1 55 6159 2296
Skype: ecorfan-mexico.s.c.
E-mail: contacto@ecorfan.org
Facebook: ECORFAN-México S. C.

Twitter: @EcorfanC

www.ecorfan.org

Holdings

Bolivia	Honduras	China	Nicaragua
Cameroon	Guatemala	France	Republic of the Congo
El Salvador	Colombia	Ecuador	Dominica
Peru	Spain	Cuba	Haití
Argentina	Paraguay	Costa Rica	Venezuela
Czech Republic			



Introducción

En el ámbito empresarial la toma de decisiones es de gran importancia y la predicción es un mecanismo de apoyo ya que es conveniente contar con una visión de lo que puede suceder en el futuro realizando un análisis sobre los datos anteriores y actuales.



**Congreso Interdisciplinario de Energías Renovables,
Mantenimiento Industrial, Mecatrónica e Informática**

2016





Minería de Datos en series de tiempo

- Técnica supervisada que se refiere al análisis de grandes conjuntos de datos secuenciales.
- Una serie tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí.
- La minería de datos en series temporales describe:
 - ❖ Si los datos presentan forma creciente, es decir tendencia.
 - ❖ Si existe influencia de ciertos periodos de cualquier unidad de tiempo, es decir la estacionalidad.
 - ❖ Si aparecen outliers (observaciones extrañas o discordantes).



Metodología ARIMA

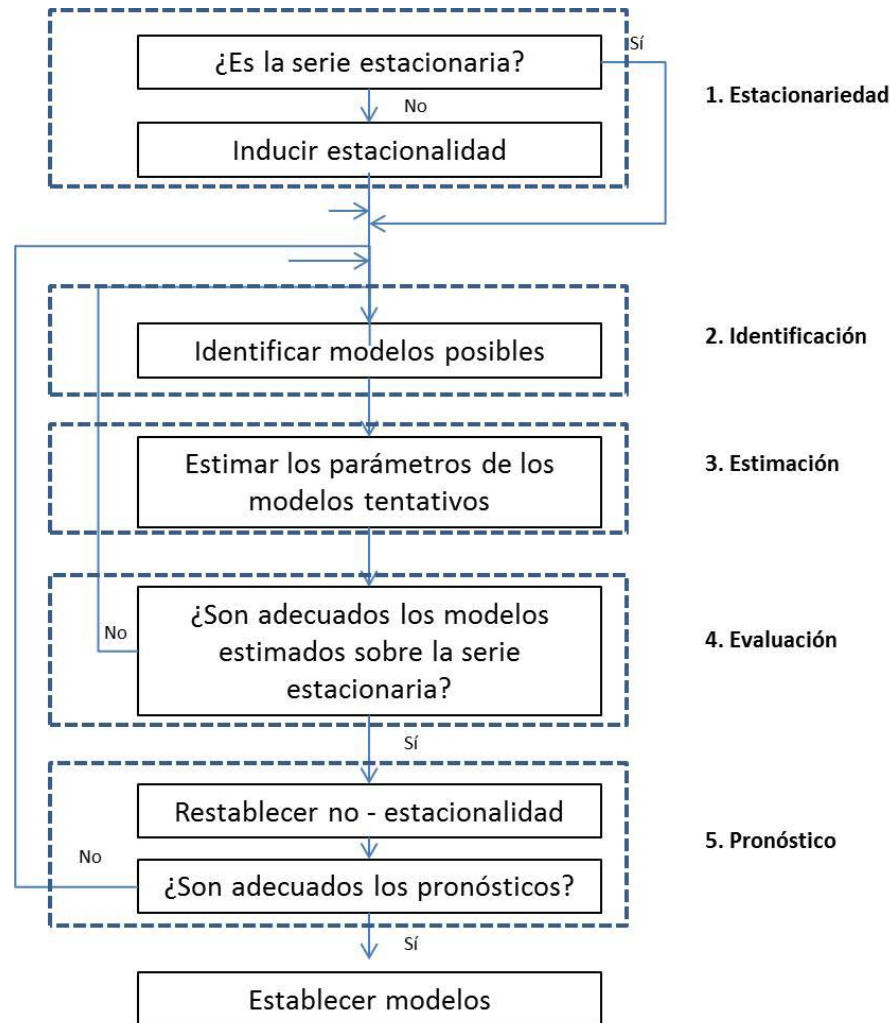
- Box y Jenkins en 1970 desarrollaron los modelos ARIMA destinados a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental.
- Se utilizan para generar pronósticos con base en una síntesis de los patrones históricos en los datos.



Metodología ARIMA

- ARIMA es la combinación de 2 modelos:
 - ❖ AR – modelos autor regresivos.
 - ❖ MA – modelos de Medias móviles
 - ❖ I - Integración de los 2 modelos

Pasos de la metodología ARIMA

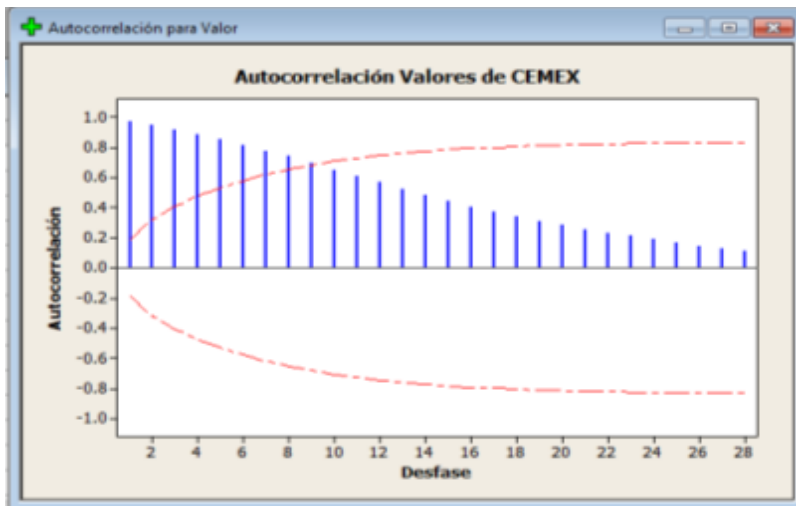


Caso de estudio aplicando metodología ARIMA

Fecha	Valor	Fecha	Valor	Fecha	Valor	Fecha	Valor
25/09/2015	11,840	05/11/2015	10,690	16/12/2015	9,100	27/01/2016	7,600
28/09/2015	11,420	06/11/2015	10,460	17/12/2015	9,140	28/01/2016	7,800
29/09/2015	11,560	09/11/2015	10,030	18/12/2015	8,980	29/01/2016	8,200
30/09/2015	11,790	10/11/2015	10,010	21/12/2015	8,830	02/02/2016	8,030
01/10/2015	11,880	11/11/2015	9,910	22/12/2015	9,090	03/02/2016	8,400
02/10/2015	11,960	12/11/2015	9,880	23/12/2015	9,740	04/02/2016	9,420
05/10/2015	12,240	13/11/2015	9,350	24/12/2015	9,720	05/02/2016	9,250
06/10/2015	12,320	17/11/2015	9,890	28/12/2015	9,390	08/02/2016	8,560
07/10/2015	12,840	18/11/2015	9,950	29/12/2015	9,310	09/02/2016	8,310
08/10/2015	13,000	19/11/2015	10,000	30/12/2015	9,260	10/02/2016	8,300
09/10/2015	12,910	20/11/2015	10,230	31/12/2015	9,430	11/02/2016	8,170
12/10/2015	12,740	23/11/2015	10,050	04/01/2016	9,120	12/02/2016	8,470
13/10/2015	12,810	24/11/2015	10,120	05/01/2016	8,990	15/02/2016	8,740
14/10/2015	13,030	25/11/2015	10,250	06/01/2016	8,970	16/02/2016	8,790
15/10/2015	12,730	26/11/2015	10,370	07/01/2016	8,260	17/02/2016	9,570
16/10/2015	12,920	27/11/2015	10,390	08/01/2016	7,770	18/02/2016	9,210
19/10/2015	12,810	30/11/2015	10,510	11/01/2016	7,830	19/02/2016	9,260
20/10/2015	12,830	01/12/2015	10,340	12/01/2016	7,990	22/02/2016	9,800
21/10/2015	12,570	02/12/2015	10,150	13/01/2016	7,720	23/02/2016	9,680
22/10/2015	11,630	03/12/2015	10,010	14/01/2016	7,960	24/02/2016	9,720
23/10/2015	11,610	04/12/2015	9,970	15/01/2016	7,920	25/02/2016	9,820
26/10/2015	11,220	07/12/2015	9,740	18/01/2016	7,850	26/02/2016	9,910
27/10/2015	10,850	08/12/2015	9,430	19/01/2016	7,520	29/02/2016	10,000
28/10/2015	11,010	09/12/2015	9,550	20/01/2016	7,370	01/03/2016	10,420
29/10/2015	10,800	10/12/2015	9,520	21/01/2016	7,880	02/03/2016	10,410
30/10/2015	10,470	11/12/2015	9,020	22/01/2016	8,050	03/03/2016	10,470
03/11/2015	10,520	14/12/2015	8,740	25/01/2016	7,470		
04/11/2015	10,620	15/12/2015	8,950	26/01/2016	7,630		

1 - Fase Estacionariedad

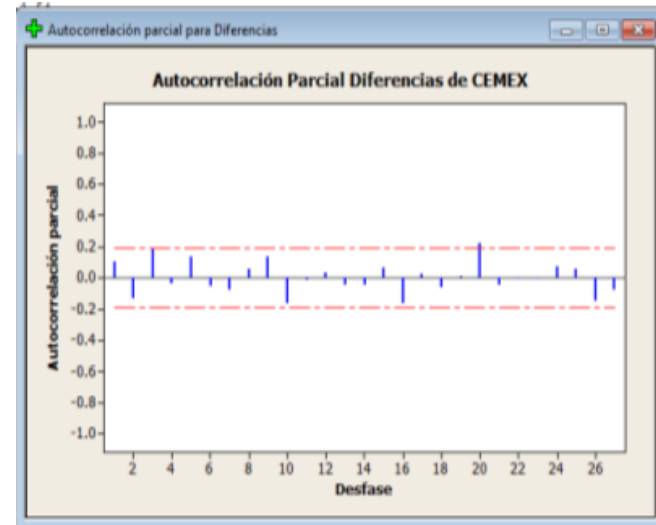
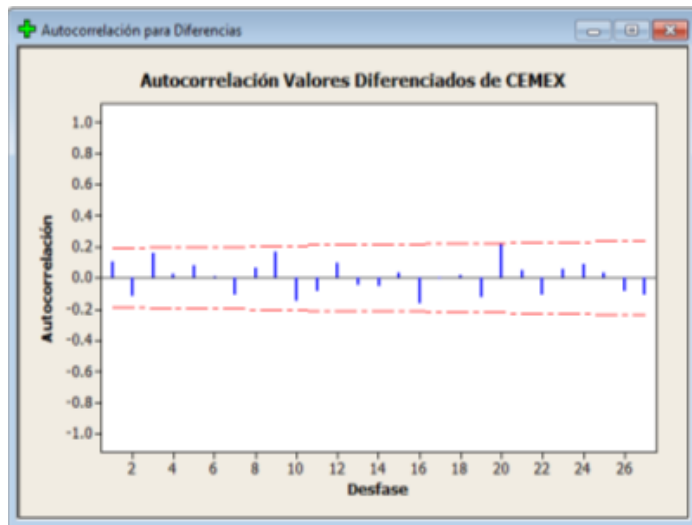
- La estacionalidad es un comportamiento o patrón en una serie de tiempo; consiste en subidas y bajadas periódicas que se presentan en forma regular en la serie de tiempo.



Serie NO estacionaria, tiene tendencia ya que los valores descienden con rapidez hasta llegar a 0.

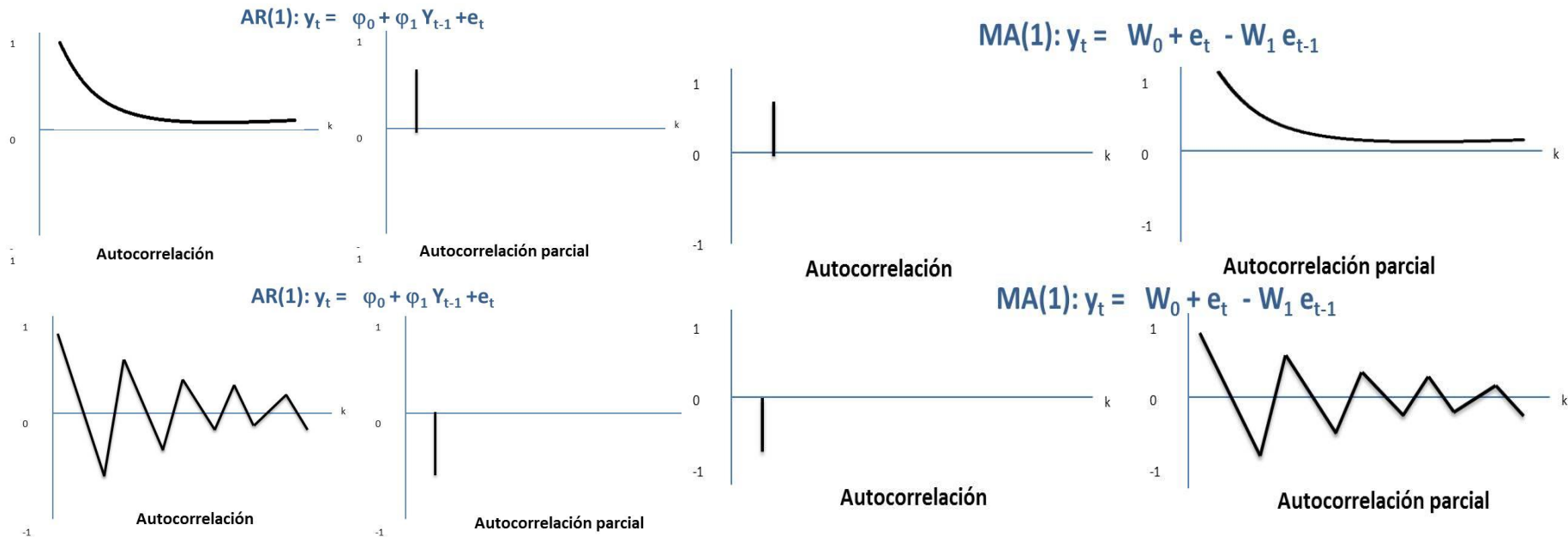
Convertir a estacionaria la serie

- Se analizan las gráficas de auto correlación y auto correlación parcial.
- Se aplica el método de diferenciación las veces que sean necesarias.



2 - Fase de Identificación

- Después de convertir la serie a estacionaria se identifica que modelo teórico es el representativo para aplicar el modelo ARIMA.



3 - Fase de Estimación

- Se aplica el modelo ARIMA con los parámetros identificados de acuerdo a las gráficas teóricas.
- Se determina un posible modelo ARIMA, formado por los elementos de p , q y d . Ejemplo:

$$\text{ARIMA } (p=1, d=1, q=0) \rightarrow \text{ARIMA}(1, 1, 0)$$

- a) p el coeficiente del modelo autorregresivo.
- b) d es el valor de la diferenciación aplicada.
- c) q el coeficiente del modelo medias móviles.

Fase de Estimación

- Analizar si los coeficientes de auto correlación se deben excluir del modelo. Se utiliza el error estándar y el valor t asociado a cada uno de los coeficientes del modelo. El error estándar tiene la siguiente fórmula:

$$t = \frac{\text{coeficiente calculado}}{\text{error estándar asociado al coeficiente calculado}}$$

Fase de Estimación

Sesión

Estimados finales de los parámetros

Tipo	Coef	SE Coef	T	P
AR 1	0.8375	0.3077	2.72	0.008
MA 1	0.7383	0.3695	2.00	0.050
Constante	0.06688	0.02371	2.82	0.006

Diferenciación: 1 Diferencia regular
 Número de observaciones: Serie original 72, después de diferenciar 71
 Residuos: SC = 37.1359 (se excluyeron pronósticos retrospectivos)
 CM = 0.5461 GL = 68

Mayor a 2 en valor absoluto,
se acepta el modelo

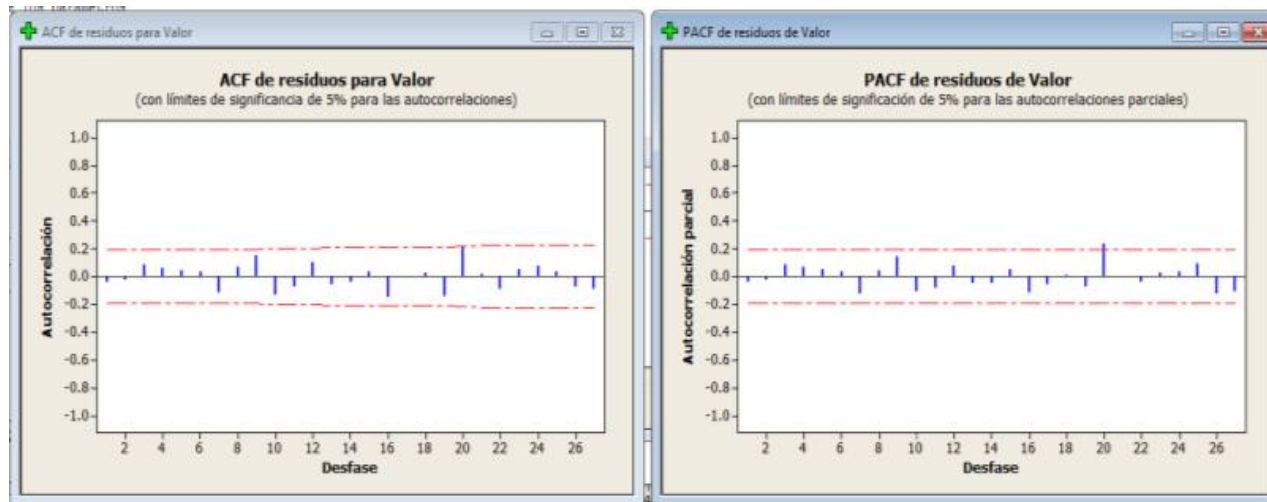
4 - Fase de Evaluación

- En esta fase se evalúa la adecuación de los modelos. Si las estimaciones de los coeficientes del modelo son significativas y cumplen las condiciones de estacionariedad e invertibilidad.

Modelo	Condiciones de estacionariedad	Condiciones de inversibilidad
AR(1)	$ \phi_1 < 1$	Ninguna
AR(2)	$\phi_1 + \phi_2 < 1$ $\phi_1 - \phi_2 < 1$ $ \phi_2 < 1$	Ninguna
MA(1)	Ninguna	$ w_1 < 1$
MA(2)	Ninguna	$w_1 + w_2 < 1$ $w_1 - w_2 < 1$ $ w_2 < 1$
ARMA(1,1)	$ \phi_1 < 1$	$ w_1 < 1$

Fase de Evaluación

Se comprueba si el modelo es estadísticamente adecuado, por lo que se analizan los residuales, a través de las funciones de auto correlación



Fase de Evaluación

- También se verifica con los resultados del valor P (probabilidad) de la estadística de Chi-cuadrada.
- Se verifica que se cumplan con un 95% de confianza.

Sesión

Estadística Chi-cuadrada modificada de Box-Pierce (Ljung-Box)

Desfase	12	24	36	48
Chi-cuadrada	10.6	25.9	38.8	46.8
pL	9	21	33	45
Valor P	0.302	0.212	0.226	0.398

5 - Fase de Pronóstico

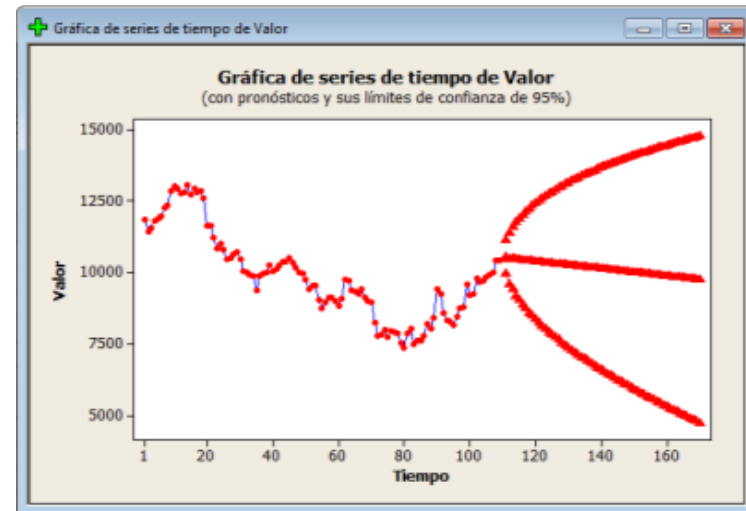
- Después de la validación se procede a generar los pronósticos de *n periodo de tiempo deseado* a partir de los datos propuestos.

Minitab - Sin título - [Sesión]

Archivo Editor Datos Calc Estadísticas Gráfica Editor Herram

Pronósticos del periodo 110

Periodo	Pronóstico	Límites 95%		Actual
		Inferior	Superior	
111	10534.4	9962.1	11106.6	
112	10470.1	9569.1	11371.1	
113	10491.6	9401.5	11591.6	
114	10455.9	9177.2	11794.6	
115	10458.3	9032.1	11884.5	
116	10435.3	8865.4	12005.3	
117	10429.3	8733.9	12124.7	
118	10412.0	8596.0	12228.0	
119	10402.1	8475.5	12328.8	
120	10387.3	8354.5	12420.2	
121	10375.8	8243.1	12508.6	
122	10362.1	8133.3	12591.0	
123	10349.9	8029.4	12670.4	
124	10336.7	7927.7	12745.7	
125	10324.1	7829.9	12818.3	
126	10311.1	7734.5	12887.7	
127	10298.4	7642.0	12954.9	
128	10285.5	7551.5	13019.5	
129	10272.7	7463.4	13082.1	
130	10259.9	7377.1	13142.7	
131	10247.1	7292.7	13201.5	
132	10234.3	7210.0	13258.6	
133	10221.4	7128.8	13314.1	
134	10208.6	7049.2	13368.1	
135	10195.8	6970.9	13420.7	
136	10183.0	6893.9	13472.0	
137	10170.2	6818.1	13522.2	
138	10157.3	6743.6	13571.1	
139	10144.5	6670.1	13619.0	
140	10131.7	6597.6	13665.8	
141	10118.9	6526.2	13711.6	
142	10106.0	6455.6	13756.5	



Conclusiones

- La metodología ARIMA permite realizar predicciones de forma fácil y eficientes para las series de datos temporales.
- La metodología ARIMA es posible aplicarse a la predicción de series temporales de diferentes indicadores productivos.
- Utilizando el software Minitab es posible realizar el proceso complejo de modelar una serie de tiempo con el propósito de pronosticar nuevos datos dentro de un periodo de corto, mediano o largo plazo, de forma fácil y eficiente.



ECORFAN®

© ECORFAN-Mexico, S.C.

No part of this document covered by the Federal Copyright Law may be reproduced, transmitted or used in any form or medium, whether graphic, electronic or mechanical, including but not limited to the following: Citations in articles and comments Bibliographical, compilation of radio or electronic journalistic data. For the effects of articles 13, 162,163 fraction I, 164 fraction I, 168, 169,209 fraction III and other relative of the Federal Law of Copyright. Violations: Be forced to prosecute under Mexican copyright law. The use of general descriptive names, registered names, trademarks, in this publication do not imply, uniformly in the absence of a specific statement, that such names are exempt from the relevant protector in laws and regulations of Mexico and therefore free for General use of the international scientific community. BCIERMIMI is part of the media of ECORFAN-Mexico, S.C., E: 94-443.F: 008- (www.ecorfan.org/ booklets)