

TRABAJO ESPECIAL

**ANALISIS DE SERIES DE TIEMPO POR AUTOCORRELACION
APLICADO A CAUDALES MEDIOS DIARIOS**

**PRESENTADO ANTE LA ILUSTRE
UNIVERSIDAD CENTRAL DE VENEZUELA
POR LOS BACHILLERES:
ANGEL D. RAMIREZ J.
VALMORE J. MORALES G.
PARA OPTAR AL TITULO DE
INGENIERO HIDROMETEOROLOGISTA**

CARACAS, ENERO 1973

TRABAJO ESPECIAL

**ANALISIS DE SERIES DE TIEMPO POR AUTOCORRELACION
APLICADO A CAUDALES MEDIOS DIARIOS**

**PRESENTADO ANTE LA ILUSTRE
UNIVERSIDAD CENTRAL DE VENEZUELA
POR LOS BACHILLERES:
ANGEL D. RAMIREZ J.
VALMORE J. MORALES G.
PARA OPTAR AL TITULO DE
INGENIERO HIDROMETEOROLOGISTA**

CARACAS, ENERO 1973

ANALISIS DE SERIES DE TIEMPO POR AUTOCORRELACION
APLICADO A CAUDALES MEDIOS DIARIOS

ALUMNOS:

Angel D. Ramírez J.

Valmore J. Morales G.

PROFESOR GUIA:



Angel Guerra Y.

A mis Padres

A mis Hermanos

Angel D. Ramírez J.

DEDICATORIA

A mis Padres

Valmore J. Morales G.

RECONOCIMIENTO

Queremos expresar nuestro agradecimiento al Profesor Guía ANGEL GUERRA, por su colaboración en la realización de este trabajo.

Al Ing. JOSE MIGUEL PEREZ GODOY, nuestro sincero reconocimiento por su asesoramiento y su valiosa cooperación que nos prestó en todo momento.

Asimismo queremos expresar nuestro agradecimiento a las siguientes dependencias :

División de Hidrología - D.G.R.H. - M.O.P.

Unidad Regional IV - División de Proyecto - D.G.R.H. - M.O.P.

Planificación de Cuencas - Oficina de Planeamiento - D.G.R.H. -
M.O.P.

Centro de Computación - Facultad de Ingeniería - U.C.V.

RESUMEN.

El propósito del presente trabajo especial, es analizar las series de tiempo de caudales medios diarios, cuyas características serán objeto de estudio.

Dicho estudio consiste en separar en las series citadas el componente determinístico o periódico, del componente estocástico o aleatorio, y reconstruir el proceso fundamental en términos de un modelo matemático.

El componente periódico es representado por la serie de Fourier, a partir de la serie original; de esta forma se obtiene la ecuación matemática del componente cíclico.

El componente estocástico consta de una parte dependiente y otra independiente, la parte dependiente fué analizada por el método de autocorrelación, definiendo un modelo autoregresivo para cada serie en particular, que representa dicha parte.

La técnica estadística de autocorrelación determina, la dependencia lineal de valores sucesivos de la serie en cuestión.

La parte independiente es completamente aleatoria.

Una vez representado el componente periódico y analizado el componente estocástico, tenemos el modelo matemático:

$$X_t = P_t + e_t$$

donde: X_t = Serie.

P_t = Componente periódico.

e_t = Componente estocástico.

a partir del cual y usando las técnicas de generación, podemos producir una serie de mayor longitud que la serie original, y que conserve sus parámetros estadísticos.

<u>INDICE DEL CONTENIDO</u>		<u>PAGINA</u>
INTRODUCCION		1
CAPITULO I METODOLOGIA		3
1-1 MODELO GENERAL		5
1-2 COMPONENTE PERIODICO		6
1-3 COMPONENTE ESTOCASTICO		9
1-3-1 Modelo Autoregresivo de primer orden.....		11
1-3-2 Modelo Autoregresivo de segundo orden.....		12
1-3-3 Modelo Autoregresivo de tercer orden.....		14
1-4 GENERACION.....		16
CAPITULO II RESULTADOS		19
CAPITULO III CONCLUSIONES		31
BIBLIOGRAFIA		33
ANEXO (GRAFICOS)		

INTRODUCCION

En todo estudio que conlleve al aprovechamiento de los recursos hidráulicos, es indispensable conocer la disponibilidad de agua. Debido al constante crecimiento de su demanda, se hace necesario, estimar la futura suplencia de ella.

Los registros de los caudales medios diarios de un río, nos dan un camino en términos de una serie de tiempo, a través del cual podemos estimar dicha disponibilidad.

La escorrentía de una cuenca es la combinación de efectos de variables que pueden ser de naturaleza determinística o estocástica.

El objeto del presente trabajo es investigar la estructura de la serie de tiempo de caudales medios diarios, separar el componente determinístico del estocástico y reconstruir el proceso fundamental en términos de un modelo matemático, el cual adecuadamente describirá el mecanismo que genera el proceso.

CAPITULO I

M E T O D O L O G I A

METODOLOGIA

Al hacer un estudio donde estén involucradas variables hidrológicas, es necesario hacer una distinción entre los valores observados, y los valores reales. Los valores reales nunca son conocidos, debido a que los datos son obtenidos afectados de errores inevitables en sus observaciones. Los valores observados tienen errores aleatorios y sistemáticos, en este caso se dice que los datos son inconsistentes (inconsistencia es la diferencia, entre el valor observado con el inherente error sistemático y el valor real). Los cambios en los principales ambientes hidrológicos ocurren por accidentes, por evoluciones lentas - en la naturaleza o por actividades humanas. Los datos afectados por esos cambios ambientales se dice que son no homogéneos (no homogéneo es la diferencia, entre el valor real y el valor virgen), hablándose de no estacionaridad en el caso de una serie de tiempo.

Un conjunto de datos históricos, observados y registrados en el tiempo, o una serie de valores en una sucesión de intervalos de tiempo, es lo que se conoce con el nombre de SERIE DE TIEMPO.

Las principales series hidrológicas son series de tiempo, un ejemplo de ello es el registro de caudales medios diarios de un río.

Las cuencas de los ríos están sometidas constantemente a cambios, debidos a la naturaleza misma y a cambios debidos a la actividad humana. Los cambios debidos a la naturaleza, son difíciles de percibir en un período corto de tiempo, ellos operan en el denominado tiempo geológico.

La serie de tiempo se dice que es estacionaria, si sus parámetros no varían con el tiempo, en caso contrario se dice que es no estacionaria.

Las series hidrológicas son generalmente no estacionarias. Sin embargo, en hidrología se toma las series como estacionarias pues se calcula un promedio

con respecto al tiempo de un solo registro y se usa luego con un sentido probabilístico.

Se hace necesario investigar cuan correcta es la suposición más general de estacionaridad.

Las causas de la no estacionaridad de una serie, son los cambios repentinos en la naturaleza, tales como terremotos, deslizamientos de tierra, cambios hechos por el hombre, etc., por ejemplo, si un río es represado, la varianza de los gastos aguas abajo de la presa sufre un cambio respecto a la media original, en otras palabras el régimen de registros del río varía y por ende sus parámetros estadísticos; también los incendios forestales pueden producir cambios repentinos en varios de sus parámetros.

En este trabajo se supone que las series son estacionarias.

1-1 MODELO GENERAL.

Una serie de tiempo está compuesta de dos elementos, uno determinístico o cíclico y otro estocástico o aleatorio; donde estos elementos, si ambos están presentes, se encuentran combinados de una manera lineal o no lineal.

El modelo general viene dado por:

$$X_t = P_t + e_t \quad (1-1)$$

P_t = Componente cíclico o periódico.

e_t = Componente estocástico.

El componente determinístico es definido por una función periódica, esa función diferente de la función aleatoria, está desprovista de la noción de probabilidad y los valores que ella asume dependen básicamente, en este caso de la variable tiempo.

Los gastos medios diarios tienden a seguir un ciclo medio anual, ellos son altos durante la estación húmeda y bajos durante la estación seca.

Los gastos medios de un día específico, tienden a variar de un año a otro alrededor de su media, estas medias diarias definen el ciclo anual, el cual es la función periódica. Este componente periódico será descrito por serie de Fourier.

El componente estocástico e_t , está gobernado por leyes de probabilidad, y consta de dos partes, una dependiente y otra independiente. La parte dependiente será analizada por métodos estadísticos, siendo la parte independiente netamente aleatoria.

1-2 COMPONENTE PERIODICO.

Una de las características más notables de los fenómenos hidrológicos, es que sus observaciones producen resultados que varían de una manera más o menos regular; tal regularidad se debe a la ocurrencia periódica de eventos astronómicos, algunos de los cuales ejercen un control directo sobre esos fenómenos, por ejemplo las fases de la luna están directamente asociadas con las mareas, los movimientos de la tierra producen cambios estacionales, etc.

Basándose en todo lo anterior, es posible representar parte de los fenómenos hidrológicos a través de una función periódica, la propiedad básica de tal función, es que se repite a si misma en un intervalo regular L , siendo L el período.

Todas las series de tiempo hidrológicas de valores diarios tales como precipitación, gastos de un río etc., tienen componentes determinísticos, los cuales pueden ser reproducidos mediante series periódicas.

Las series hidrológicas tienen un ciclo básico, determinado por los fenómenos naturales el cual es casi periódico, para las series de valores diarios ese ciclo es de 365 días.

El componente periódico, P_t , del modelo general fue representado por la serie de Fourier, de la siguiente forma:

$$P_t = \bar{X} + \sum_{j=1}^m (A_j \cos \frac{2\pi j}{L} t + B_j \sin \frac{2\pi j}{L} t) \quad (1-2)$$

donde: $\bar{X}$ = Media de la serie de datos

m = N° de armónicos.

L = Período.

$$A_j = \frac{2}{n} \sum_{t=1}^n x_t \cos \frac{2\pi j}{L} t \quad (1-3)$$

$$B_j = \frac{2}{n} \sum_{t=1}^n x_t \sin \frac{2\pi j}{L} t \quad (1-4)$$

donde: $n = N^\circ$ de valores observados en la serie o número de datos.

El término $C^2 = A^2 + B^2$ es lo que se denomina amplitud de cada armónico.

La significancia de un armónico se mide por la varianza del armónico expresada en porcentaje respecto a la varianza de la serie.

El parámetro usado para la prueba de significancia de los armónicos, es la varianza de cada armónico, la cual es dada por los coeficientes de Fourier y se expresa de la siguiente forma :

$$\frac{C_j^2}{2} = \frac{1}{2} (A_j^2 + B_j^2) \quad (1-5)$$

El primer test que prueba la significancia de los armónicos fué desarrollado por Schuster, posteriormente Walker y Fischer desarrollaron otros test.

El objetivo que se persigue al emplear un test que pruebe la significancia de los armónicos es la utilización del menor número de parámetros.

En este trabajo se emplea el test de Fischer que prueba la significancia usando la relación :

$$g_1 = \frac{C_{\max}^2/2}{\sigma_x^2} \quad (1-6)$$

donde: $C_{\max}^2$ = Es el valor mayor de la secuencia de C_j^2 .

σ_x^2 = Varianza de la serie de datos.

Esto para el caso de un armónico.

Un armónico se dice que es significativo si $g_1 > g_c$ con un nivel de probabilidad seleccionado, este nivel de probabilidad de que el valor g_1 exceda al valor crítico g_c viene dado por:

$$P_f = m (1 - g_c)^{m-1} \quad (1-7)$$

dónde: $m = N/2$ siendo N la longitud de la serie.

g_c = Valor crítico.

P_f = Probabilidad de que g_1 exceda al valor g_c .

Despejando g_c tenemos:

$$g_c = 1 - \sqrt[m-1]{P_f/m} \quad (1-8)$$

En el caso de que dos o más armónicos sean significativos entonces se hace $C_{\max}^2 = C_1^2$ y si se halla ser significativo con $g_1 > g_c$ para un nivel de probabilidad seleccionado se calcula y se sustrae de la serie entonces el próximo valor C_2^2 es probado contra $\sigma_x^2 - C_1^2/2$ tal que el nuevo valor g_2 es :

$$g_2 = \frac{C_2^2/2}{\sigma_x^2 - C_1^2/2} \quad (1-9)$$

y así para los sucesivos armónicos. En general tendremos que :

$$g_i = \frac{C_i^2/2}{\sigma_x^2 - \sum_{j=1}^{i-1} C_j^2/2} \quad (1-10)$$

En donde los $C_j^2/2$ de la sumatoria son mayores que $C_i^2/2$, o sea que ésta representa valores $C_j > C_i$.

1-3 COMPONENTE ESTOCASTICO.

Asumiendo que el componente periódico, P_t , ha sido encontrado y separado de la serie, los residuos e_t constituyen la variable estocástica a analizar.

$$e_t = X_t - P_t \quad (1-11)$$

Un proceso estocástico es la abstracción matemática de un proceso empírico, el cual es gobernado por leyes de probabilidad. Este proceso se dice que es linealmente dependiente si sus valores sucesivos están relacionados (o se pueden representar) por un modelo matemático lineal.

La variable estocástica e_t consta de una parte dependiente y otra independiente.

La parte dependiente se analiza por autocorrelación, lo cual determina la dependencia lineal entre los valores sucesivos de la serie.

El coeficiente de autocorrelación es definido por la expresión :

$$r_k = \frac{\text{Cov}(e_i, e_i + k)}{(\text{Var}e_i \cdot \text{Var}e_i + k)^{1/2}} \quad (1-12)$$

La cual es equivalente a :

$$r_k = \frac{\frac{1}{n-k} \sum_{t=1}^{n-k} e_t e_{t+k} - \frac{1}{(n-k)^2} \left(\sum_{t=1}^{n-k} e_t \right) \left(\sum_{t=1}^{n-k} e_{t+k} \right)}{\left[\frac{1}{n-k} \sum_{t=1}^{n-k} e_t^2 - \frac{1}{(n-k)^2} \left(\sum_{t=1}^{n-k} e_t \right)^2 \right]^{1/2} \left[\frac{1}{n-k} \sum_{t=1}^{n-k} e_{t+k}^2 - \frac{1}{(n-k)^2} \left(\sum_{t=1}^{n-k} e_{t+k} \right)^2 \right]^{1/2}} \quad (1-13)$$

La función $r_k = f(k)$ es llamada función de autocorrelación o simplemente "Correlograma", el cual representa propiedades relativas a la independencia o dependencia de la serie. El uso de los "Correlogramas" para la investigación de series hidrológicas se basa en analogías con "Correlogramas" de procesos estocásticos conocidos. Para la aplicación de esta analogía es necesario conocer los "Correlogramas" de muchos tipos de series.

Por inferencia estadística se puede determinar cuan aproximado es el "Correlograma" de una serie hidrológica a uno de un proceso conocido. Un procedimiento analítico objetivo tal como la inferencia cuantitativa sería empleado para aceptar o rechazar esa hipótesis, por lo tanto se hace necesario conocer la forma y propiedades de los "Correlogramas" de procesos estocásticos conocidos, interpretando esos "Correlogramas".

El análisis de los "Correlogramas" se hace de la siguiente forma :

- a. Se selecciona un modelo matemático hipotético de series dependientes de un número dado de parámetros, determinado por una familia infinita de curvas.
- b. Se estiman los parámetros, lo cual es equivalente a seleccionar una curva de esa familia.

Dependiendo de la exactitud del ajuste de esta curva al "Correlograma" de la serie observada, el modelo matemático es aceptado o rechazado. Si es rechazado se selecciona un nuevo modelo y se repite el procedimiento.

La representación de la dependencia lineal de los procesos estocásticos se hace mediante los modelos autoregresivos.

En un modelo autoregresivo el valor de una observación se asume que depende de observaciones pasadas. En la práctica se observa que el efecto de las observaciones precedentes decrece a medida que transcurre el tiempo.

La expresión matemática del modelo general autoregresivo de orden m , también llamado modelo de Markov de orden m es :

$$e_t = a_1 e_{t-1} + a_2 e_{t-2} + \dots + a_m e_{t-m} + n_t \quad (1-14)$$

donde: e_t = Magnitud del evento en consideración en el tiempo t .

a_m = Parámetros autoregresivos obtenidos por regresión múltiple lineal.

n_t = Variable aleatoria independiente (parte independiente)

El "Correlograma" de este modelo general viene expresado por :

$$r_k = a_1 r_{k-1} + a_2 r_{k-2} + \dots + a_m r_{k-m} \quad (1-15).$$

dónde los r_k son estimados de la muestra k = intervalo de tiempo.

1-3-1 Modelo autoregresivo de primer orden.

Este modelo autoregresivo también llamado primer modelo de Markov, asume que los valores de un evento en un cierto tiempo t , son solo dependientes del valor del evento precedente, en este caso $m = 1$ en el modelo general autoregresivo y de la ecuación (1-14) tenemos :

$$e_t = a_1 e_{t-1} + n_t \quad (1-16)$$

donde: $a_1 = r_1$ = Primer coeficiente de autocorrelación.

n_t = Variable aleatoria normalizada $E(n) = 0$.

De la ecuación (1-15) obtenemos el "Correlograma" para este primer modelo, haciendo $m = 1$

$$r_k = r_1^k \quad (1-17)$$

Este correlograma decrece exponencialmente a cero cuando $k \rightarrow \infty$.

Si el valor esperado $E(n) = 0$ siendo la varianza de n igual a σ_n^2 ; la varianza de e siendo $\text{Var } e_t = \text{Var } e_{t-1} = \sigma_e^2$ viene dada por :

$$\text{Var } e = \sigma_e^2 = \frac{\sigma_n^2}{1 - r_1^2} \quad (1-18)$$

1-3-2 Modelo autoregresivo de segundo orden.

Llamado también segundo modelo de Markov, asume que los valores de un evento en un cierto tiempo t dependen de los dos eventos precedentes.

En este caso $m = 2$ y de la ecuación (1-14) tenemos :

$$e_t = a_1 e_{t-1} + a_2 e_{t-2} + n_t \quad (1-19)$$

donde: n_t = Variable aleatoria normalizada $E(n_t) = 0$.

De la ecuación (1-15) se obtiene el "Correlograma" para este modelo :

$$r_k = a_1 r_{k-1} + a_2 r_{k-2} \quad (1-20)$$

Este "Correlograma" tiene el primero y segundo coeficiente de autocorrelación iguales a los obtenidos en la serie. Para $k \geq 2$ deben ser calculados por un proceso iterativo de la ecuación (1-20).

Los parámetros a_1 y a_2 vienen dados por las siguientes expresiones :

$$a_1 = \frac{r_1 - r_1 r_2}{1 - r_1^2} \quad (1-21)$$

$$a_2 = \frac{r_2 - r_1^2}{1 - r_1^2} \quad (1-22)$$

Si el valor esperado $E(n) = 0$ siendo la varianza de n igual a σ_n^2 ; la varianza de e siendo $\text{Var } e_t = \text{Var } e_{t-1}$ y $2 \text{Cov}(a_1 e_{i-1}, a_2 e_{i-2}) = 2 a_1 a_2 \text{Cov}(e_{i-1}, e_{i-2}) = 2 a_1 a_2 r_1$ viene dada por la expresión :

$$\text{Var } e_i = \sigma_e^2 = \frac{\sigma_n^2}{1 - a_1^2 - a_2^2 - 2 a_1 a_2 r_1} = \frac{\sigma_n^2 (1 - r_1^2)}{1 - 2 r_1^2 (1 - r_2) - r_2^2} \quad (1-23)$$

El cuadrado del coeficiente de correlación múltiple R viene dado por :

$$R^2 = \frac{\sigma_e^2 - \sigma_n^2}{\sigma_e^2} \quad (1-24)$$

Obteniéndose para este segundo modelo la siguiente expresión :

$$R^2 = \frac{r_1^2 + r_2^2 - 2r_1^2 r_2}{1 - r_1^2} \quad (1-25)$$

El cual representa la parte de la varianza de e explicada por la regresión. El valor $(1-R^2)$ representa la importancia de la parte independiente (n) , mientras que R^2 representa la importancia autoregresiva del modelo.

Para la condición de estacionaridad del "Correlograma" de este segundo modelo se debe cumplir que :

$$a_1^2 + 4a_2 \leq 0$$

1-3-3 Modelo autoregresivo de tercer orden.

Denominado también tercer modelo de Markov se obtiene = haciendo $m = 3$ en la ecuación (1-14).

$$e_t = a_1 e_{t-1} + a_2 e_{t-2} + a_3 e_{t-3} + n_t \quad (1-26)$$

n_t = Variable aleatoria normalizada $E(n) = 0$.

cuyo "Correlograma" a partir de la ecuación (1-15) resulta:

$$r_k = a_1 r_{k-1} + a_2 r_{k-2} + a_3 r_{k-3} \quad (1-27)$$

obteniéndose a_1 , a_2 y a_3 a partir de las siguientes expresiones :

$$a_1 = \frac{(1-r_1)^2 (r_1-r_3) - (1-r_2) (r_1 r_2 - r_3)}{(1-r_2) (1-2r_1^2 + r_2)} \quad (1-28)$$

$$a_2 = \frac{r_2 + r_2^2 - r_1^2 - r_1 r_3}{1 - 2r_1^2 + r_2} \quad (1-29)$$

$$a_3 = \frac{(r_1 - r_3)(r_1^2 - r_2) - (1 - r_2)(r_1 r_2 - r_3)}{(1 - r_2)(1 - 2r_1^2 + r_2)} \quad (1-30)$$

Si el valor $E(n) = 0$ siendo la varianza de n igual a σ_n^2 ; la varianza de e siendo $V_{ar} e_t = V_{ar} e_{t-1} = \sigma_e^2$ viene dada por :

$$V_{ar} e = \sigma_e^2 = \frac{\sigma_n^2}{1 - a_1^2 - a_2^2 - a_3^2 - 2a_1 a_2 r_1 - 2a_1 a_3 r_2 - 2a_2 a_3 r_1} \quad (1-31)$$

y de la ecuación (1-24) tenemos que el cuadrado del coeficiente de correlación múltiple R obtenido para este tercer modelo es:

$$R^2 = a_1^2 + a_2^2 + a_3^2 + 2a_1 a_2 r_1 + 2a_1 a_3 r_2 + 2a_2 a_3 r_1 \quad (1-32)$$

1.4 GENERACION.

En el análisis de los recursos del agua, un problema frecuente es la falta de suficientes registros hidrológicos. Los registros existentes son generalmente cortos para ser usados en la planificación y elaboración de proyectos de obras hidráulicas; en la mayoría de los casos, se encuentran problemas de estimación del futuro uso del agua, basándose en registros hidrológicos deficientes.

Las técnicas de generación de datos hidrológicos, actualmente son una ayuda para la solución de esos problemas. Esta técnica consiste en extraer la máxima información de la muestra, mediante métodos que permiten mantener las propiedades de los datos (media, varianza, etc).

Los métodos de generación de datos hidrológicos a partir de datos históricos limitados han sido llamados hidrología paramétrica e hidrología estocástica.

La hidrología paramétrica trata con relaciones determinísticas (significa que existe una relación causa efecto, es decir se obtiene una sola respuesta a cada entrada) entre los parámetros físicos envueltos en los procesos hidrológicos.

La hidrología estocástica emplea los métodos estadísticos y probabilidad en el análisis de los procesos hidrológicos, sintetizando información y suponiendo la intervención de una variable aleatoria.

En el campo de la hidrología estocástica, la generación de datos hidrológicos se basa en la estadística y técnicas matemáticas, para crear series hipotéticas de registros que estadísticamente no se di

ferencian de las series históricas.

Los métodos de generación requieren del establecimiento de un modelo que genere el proceso hidrológico. El modelo se obtiene a partir de los parámetros de la serie histórica hallados por análisis estadístico.

Representada la serie histórica por el modelo descrito en este capítulo, estamos en condiciones de generar una serie hipotética más larga que la original.

Es de aclarar que los datos generados no sustituyen a los históricos (base de la generación) sino que nos ayudan en la toma de decisiones.

CAPITULO II

R E S U L T A D O S

RESULTADOS.

En este capítulo se dan los pasos a través de los cuales se obtuvieron los resultados de este trabajo.

Los datos básicos empleados fueron los gastos medios diarios, tomados en forma continua, se seleccionaron un total de 5 ríos, los cuales se indican en la tabla N° 1.

Se recurrió a la programación electrónica, debido al volumen de datos empleados y a lo laborioso de los cálculos matemáticos, a tal efecto se elaboró un programa en lenguaje FORTRAN IV en el cual se efectúa todo el proceso matemático descrito en el capítulo I.

Los gráficos realizados en este trabajo fueron obtenidos utilizando el sistema de delineación digital del Centro de Computación de la Facultad de Ingeniería.

El registro de los caudales medios diarios (serie original), para cada río en particular puede observarse en las figuras 1, 2, 3, 4 y 5 (del anexo) y sus parámetros correspondientes en la tabla N° 2.

A la serie de tiempo de gastos medios diarios se le ajustó la serie de Fourier, obteniéndose de esta forma el componente periódico. Este ajuste se hizo con un máximo de seis armónicos y de estos se hallaron mediante el test de Fischer los significativos.

Obtenido el componente periódico (serie de Fourier ajustada a los gastos medios diarios), se separó de la serie original. Este componente puede verse representado gráficamente en las figuras 6, 7, 8, 9 y 10 (anexo) y sus parámetros correspondientes en la tabla N° 3.

Una vez separado el componente periódico de la serie original, se procedió a analizar los residuos e_t

$$e_t = X_t - P_t \quad (2-1)$$

donde P_t es el componente periódico caracterizado por los parámetros determinados a través de la serie de Fourier.

Se hallaron las desviaciones standard de los datos respecto a la serie de Fourier ajustada a los valores medios. A estas desviaciones standard, se les ajustó otra serie de Fourier, para compensar el efecto de un cambio de varianza en el transcurso del año.

Con la representación del componente periódico y de las desviaciones standard se procedió a normalizar la serie de residuos.

$$e_t = \frac{X_t - P_t}{S_t} \quad (2-2)$$

donde: S_t = Serie de desviaciones standard.

Esta normalización no produce en los residuos, e_t , media cero ni varianza uno; por lo que se hace necesario normalizar la serie, e_t , respecto a su media y desviación standard (media igual a cero y varianza igual a uno).

La serie de residuos, e_t , dada por la ecuación (2-2) constituye el componente estocástico, el cual consta de una parte dependiente y otra independiente.

La dependencia de la serie e_t , se mide por los coeficientes de autocorrelación. El "Correlograma" correspondiente a esta serie y a los modelos de Markov (modelos autoregresivos) pueden observarse gráficamente para cada

río en las figuras 11, 12, 13, 14 y 15 del anexo.

Obtenidos los coeficientes de autocorrelación, se procedió a ajustar los modelos autoregresivos seleccionados. Para representar la parte dependiente por estos modelos se hace necesario ver hasta que punto la serie es dependiente. Cuando los coeficientes de autocorrelación son iguales a cero la serie es independiente. Anderson nos da los límites de tolerancia para aceptar que los r_k son iguales a cero y vienen dados por la siguiente expresión :

$$S = \frac{-1 \pm 1.645 \sqrt{N-2}}{N-1} \quad (2-3)$$

donde: N = Número de datos.

la cual nos da los límites de confianza para un nivel del 95%.

Se ajustó el primer modelo dado por la expresión :

$$e_t = r_1 e_{t-1} + n_t \quad (2-4)$$

donde e_t y r_1 son valores conocidos, pudiéndose determinar la variable independiente n_t ; despejando de la ecuación anterior tenemos :

$$n_t = e_t - r_1 e_{t-1} \quad (2-5)$$

Se hallaron las varianzas de e_t y n_t , (σ_e^2 y σ_n^2 respectivamente) para luego probar el ajuste de este modelo. El test que nos da la validez del ajuste viene dado por la siguiente expresión.

$$V_{ar} n_t = \sigma_e^2 (1-r_1^2) \quad (2-6)$$

en donde $V_{ar} n_t$ es la varianza teórica que debe ser igual a la varianza de los n_t (σn^2) hallados mediante la ecuación (2-4).

Los resultados se pueden ver en la tabla N° 4.

De igual forma se ajustó el segundo modelo dado por :

$$e_t = a_1 e_{t-1} + a_2 e_{t-2} + n_t \quad (2-7)$$

Los valores a_1 y a_2 se calcularon a partir de las ecuaciones (1-21) y (1-22) respectivamente.

El test que nos da la validez del ajuste de este modelo viene dado por :

$$V_{ar} n_t = \sigma e^2 (1 - a_1^2 - a_2^2 - 2a_1 a_2 r_1) \quad (2-8)$$

donde $V_{ar} n_t$ es la varianza teórica, la cual debe ser igual a la calculada σn^2 , para aceptar el modelo. Los resultados pueden verse en la tabla 5.

Por último se ajustó el tercer modelo dado por la expresión :

$$e_t = a_1 e_{t-1} + a_2 e_{t-2} + a_3 e_{t-3} + n_t \quad (2-9)$$

a_1 , a_2 y a_3 fueron calculadas de las ecuaciones (1-28), (1-29) y (1-30) respectivamente.

La varianza de n_t , (σn^2) se obtuvo y se comparó con el valor de $V_{ar} n_t$ dado por :

$$V_{ar} n_t = \sigma e^2 (1 - a_1^2 - a_2^2 - a_3^2 - 2a_1 a_2 r_1 - 2a_1 a_3 r_2 - 2a_2 a_3 r_1)$$

Si σ_n^2 , es igual a la varianza teórica $V_{ar} n_t$, este modelo es aceptado. Los resultados se presentan en la tabla 6.

TABLA N° 1
RIOS SELECCIONADOS

RIO	ESTACION	ESTADO	PERIODO DE REGISTRO
Guache	Puente viejo	Portuguesa	1953 - 1966
Motatán	Agua Viva	Trujillo	1934 - 1969
Paguey	El Paso	Barinas	1954 - 1966
Portuguesa	Puente Portuguesa	Portuguesa	1958 - 1970
Torbes	Sabaneta	Táchira	1963 - 1971

TABLA N° 2

PARAMETROS DE LA SERIE DE GASTOS MEDIOS DIARIOS

RIO	Media de la serie	Varianza	Desviación Típica
Guache	13.71	244.25	15.63
Motatán	34.92	941.75	30.68
Paguey	52.03	2515.14	50.15
Portuguesa	29.24	1130.76	33.63
Torbes	7.83	48.30	6.95

TABLA N° 3

COMPONENTE PERIODICO

Número de armónicos significativos		Guache	Motatán	Paguey	Portuguesa	Torbes
1	A	- 10.395784	0.086026	- 34.436768	- 24.219360	- 3.958238
	B	- 8.041150	- 10.547888	- 29.095627	- 20.389832	- 3.319316
	Var Expl.	0.353602	0.059073	0.404042	0.443208	0.276266
2	A	4.239383	3.782023	3.869965	8.057096	1.555787
	B	- 1.419933	- 12.459227	- 4.178493	3.690544	0.028380
	Var Expl.	0.040919	0.090010	0.006448	0.034728	0.025067
3	A	- 3.938276	- 1.138659	- 1.805053	- 5.715506	- 0.685299
	B	1.919232	- 0.470031	4.887314	0.077412	- 0.354061
	Var Expl.	0.039291	0.000806	0.005396	0.014447	0.006160
4	A	1.116645	- 3.626812	- 1.624478	2.561127	
	B	0.755619	- 1.374882	- 4.564468	- 0.024320	
	Var Expl.	0.003721	0.007987	0.004666	0.002901	
5	A	- 0.404575		- 1.474357	- 1.559852	
	B	1.054930		3.836670	0.768458	
	Var Expl.	0.002613		0.003358	0.001337	
6					2.219556	
					- 0.899687	
					0.002536	
Total % Var Expl.		44.0146	15.7876	42.3910	49.9157	30.7493

TABLA N° 4
PRIMER MODELO DE MARKOV

RIO	Coeficiente de autocorrelación r_1	Varianza de los n_t	
		Calculada (σ_n^2)	Teórica ($\text{Var } n_t$)
Guache	- 0.002690	0.9999	0.9999
Motatán	0.855098	0.2691	0.2688
Paguey	- 0.116386	0.9932	0.9855
Portuguesa	0.083792	0.9935	0.9935
Torbes	0.756018	0.4284	0.4284

TABLA N° 5
SEGUNDO MODELO DE MARKOV

RIOS	a_1	a_2	R^2	Varianza de la serie de residuos	
				Calculada (σ_n^2)	Teórica ($V_{ar} n_t$)
Guache	- 0.002691	- 0.000327	0.000007	0.9999	0.9999
Motatán	0.693426	0.189069	0.740805	0.2595	0.2592
Paguey	- 0.121063	- 0.040183	0.015138	0.9848	0.9816
Portuguesa	0.082385	0.016795	0.007300	0.9932	0.9932
Torbes	0.585412	0.225664	0.593381	0.4065	0.4066

TABLA N° 6
TERCER MODELO DE MARKOV

RIOS	a_1	a_2	a_3	R^2	Varianza de la serie de Residuos	
					Calculada (σ_n^2)	Teórica ($Var n_t$)
Guache	- 0.002691	- 0.000331	0.001472	0.000010	0.9999	0.9999
Motatán	0.669876	0.102696	0.124556	0.744800	0.2554	0.2552
Paguey	- 0.121669	- 0.042010	- 0.015097	0.015363	0.9845	0.9809
Portuguesa	- 0.082072	0.015260	0.018638	0.007646	0.9929	0.9929
Torbes	0.543328	0.116491	0.186487	0.607520	0.3923	0.3925

CAPITULO III

CONCLUSIONES

CONCLUSIONES.

La metodología empleada en este trabajo supone a la serie de tiempo formada de dos partes, la determinística y la estocástica.

$$X_t = P_t + e_t$$

Por lo general en algunos casos aparece un tercer componente llamado de tendencia que puede ser lineal o no lineal, o un salto que ocurre en la serie de datos, tomando el modelo general la forma de la siguiente expresión:

$$X_t = P_t + R_t + e_t$$

R_t = Componente de tendencia.

Este componente se presenta debido a cambios repentinos hechos por el hombre o por la naturaleza misma y se refleja generalmente en la media, varianza, coeficientes de autocorrelación y otros parámetros afectando la distribución de la serie de tiempo. Este componente cuando se presenta debe ser analizado por otros métodos que no son estudiados en este trabajo.

Este caso parece ser el presentado en los ríos Motatán y Torbes ya que al analizarse el "Correlograma" de la serie de residuos para estos ríos se observa que la dependencia persiste hasta los 50 días (Motatán) y 30 días (Torbes), dependencia que en la naturaleza no se presenta más allá del cuarto día. Se puede decir que el modelo de este estudio no es aplicable al caso de estos dos ríos.

Para el caso del río Guache, analizando sus resultados se puede concluir que consta del componente periódico y estocástico, siendo el estocástico

co netamente aleatorio ya que los esquemas autoregresivos no son aplicables puesto que no presentan dependencia lineal.

En los ríos Paguey y Portuguesa aparece dependencia lineal en los resíduos (e_t), aunque muy leve, pudiéndose ajustar el primero o el segundo modelo autoregresivo.

Debido al volumen de datos que se utilizan en este tipo de estudio, a la dificultad de su obtención, procesamiento, etc., no fué posible extenderse a más ríos como hubiese sido lo ideal.

B I B L I O G R A F I A

ADAM, PETER. Sistema de Delineación Digital. Instituto Tecnológico, Facultad de Ingeniería, U.C.V., Mayo 1972.

DIXON y MASSEY. Introducción al Análisis Estadístico, 2^a Edición 1966. McGraw-Hill.

FIERING, MYRON B. and JACKSON, BARBARA B. Synthetic Stream Flows. American Geophysical Union, 1971.

MEYER, PAUL L. Introductory Probability and Statistical Applications, 2^d Edition 1970. Addison Wesley Publishing Company.

NATIONAL RESEARCH COUNCIL OF CANADA. Statistical Methods in Hydrology. Proceedings of Hydrology Symposium N^o 5. McGill University, February, 1966.

ORGANICK, ELLIOT. A Fortran IV Primer, 2^d Edition 1967. Addison-Wesley Publishing Company.

QUIMPO, RAFAEL G. Stochastic Model of Daily River Flow Sequences. Hydrology Papers. Colorado State University. Paper N^o 18, - February, 1967. Fort Collins, Colorado.

YEVJEVICH, VUJICA, Probability and Statistics in Hydrology. Water Resources Publications, 1972.

YEVJEVICH, VUJICA. Stochastic Processes in Hydrology. Water Resources Publications, 1972.

ANEXO

GRAFICOS

RIO GUACHE - EN PUENTE VIEJO 1953 - 1966
GASTOS MEDIOS DIARIOS

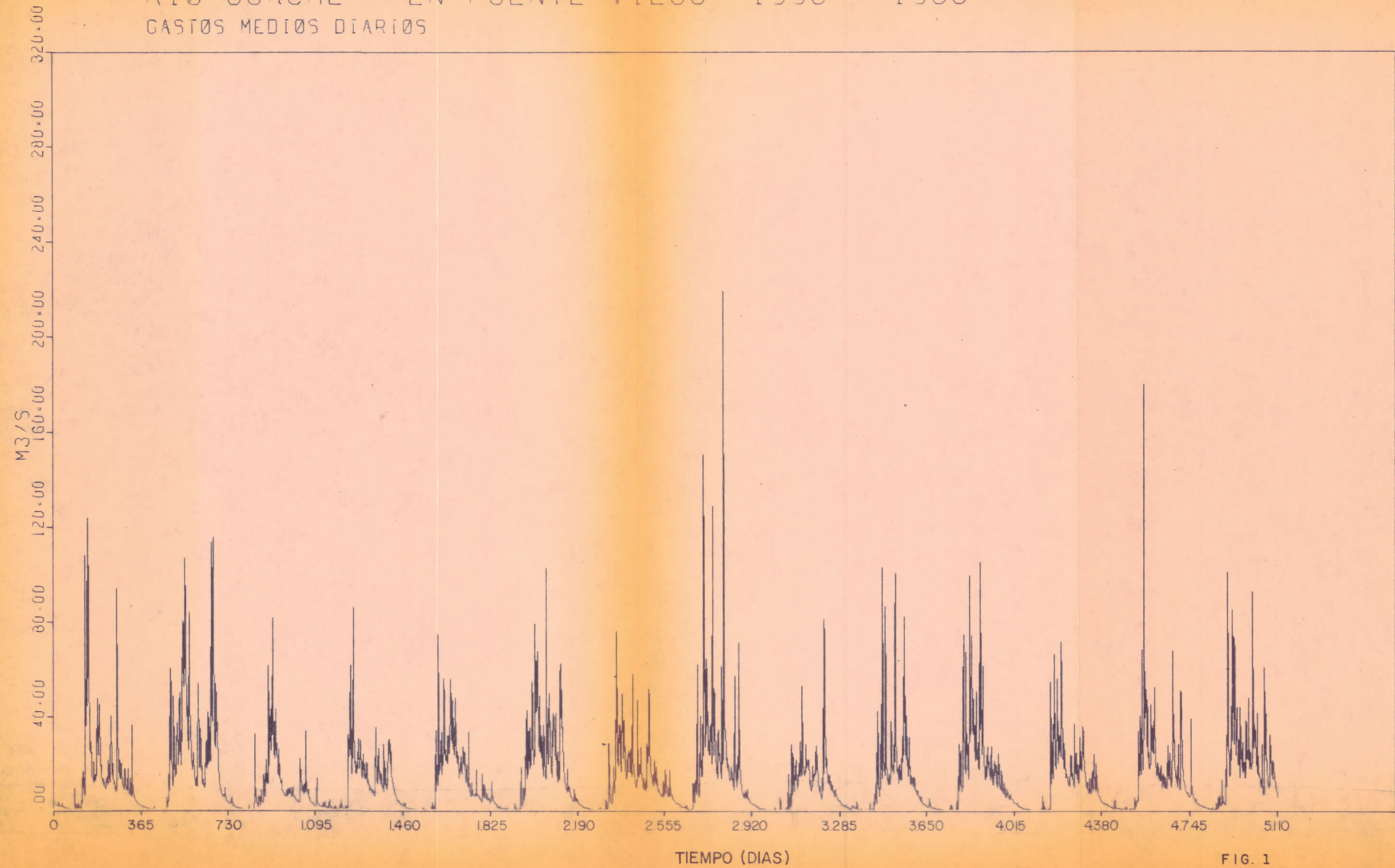


FIG. 1

RÍO MOTATÁN - EN AGUA VIVA 1943 - 1969
DATOS MEDIOS DIARIOS

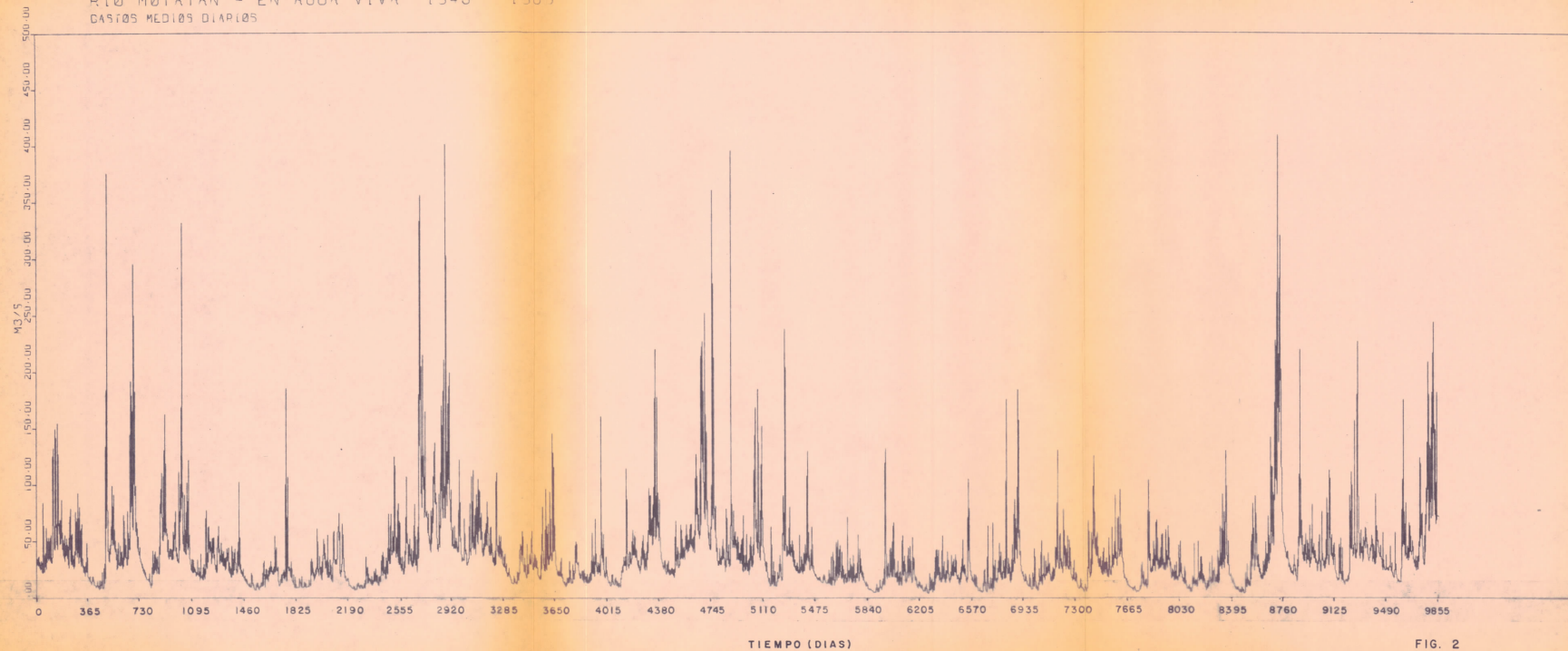


FIG. 2

RIO PAGUEY - EN EL PASO 1954 - 1966
GASTOS MEDIOS DIARIOS

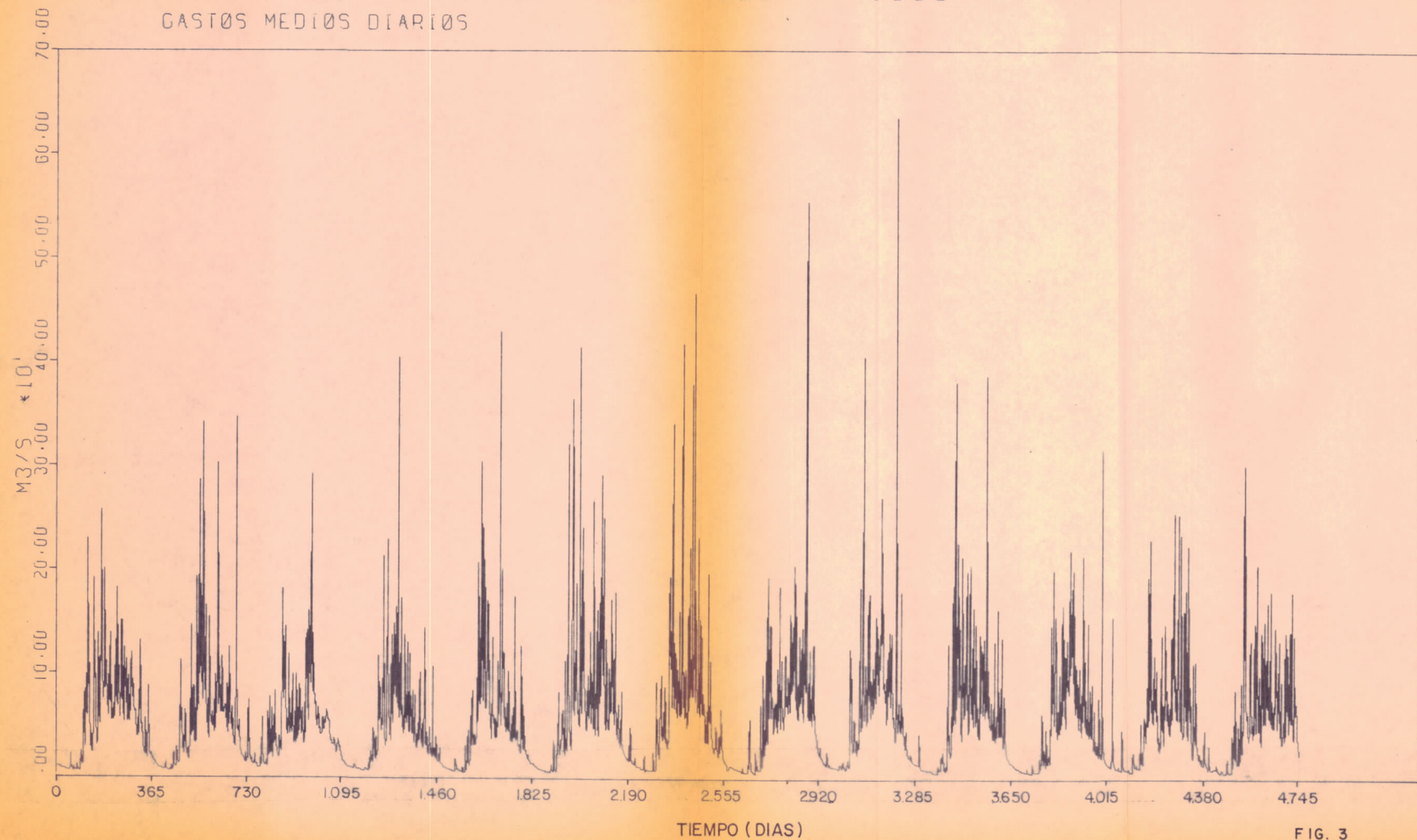


FIG. 3

RIO PORTUGUESA EN PUENTE PORTUGUESA 1958 - 1970
GASTOS MEDIOS DIARIOS

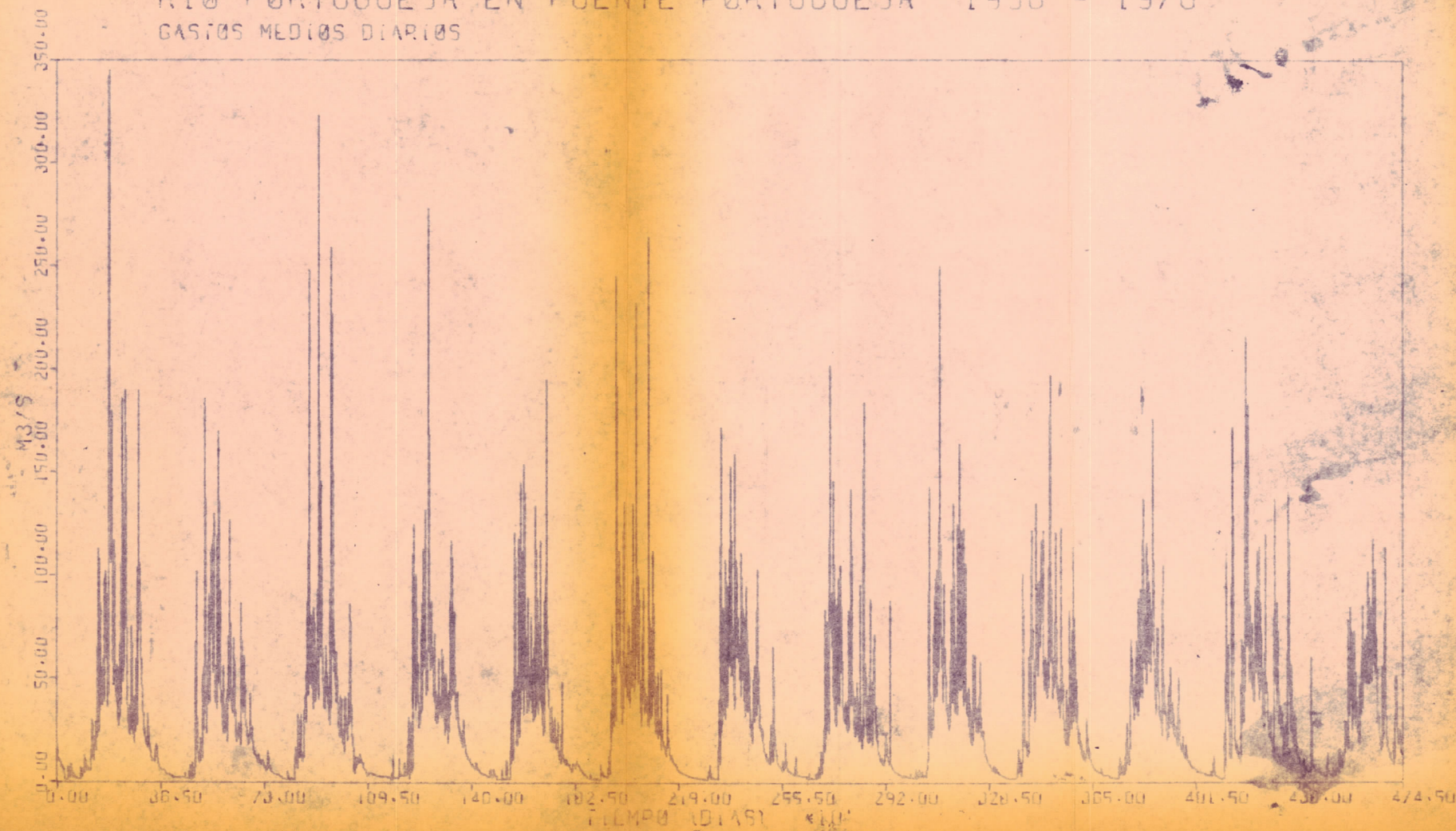


FIG. 4

RIO TØRBES - EN SABANETA 1963 - 1971
GASTOS MEDIOS DIARIOS

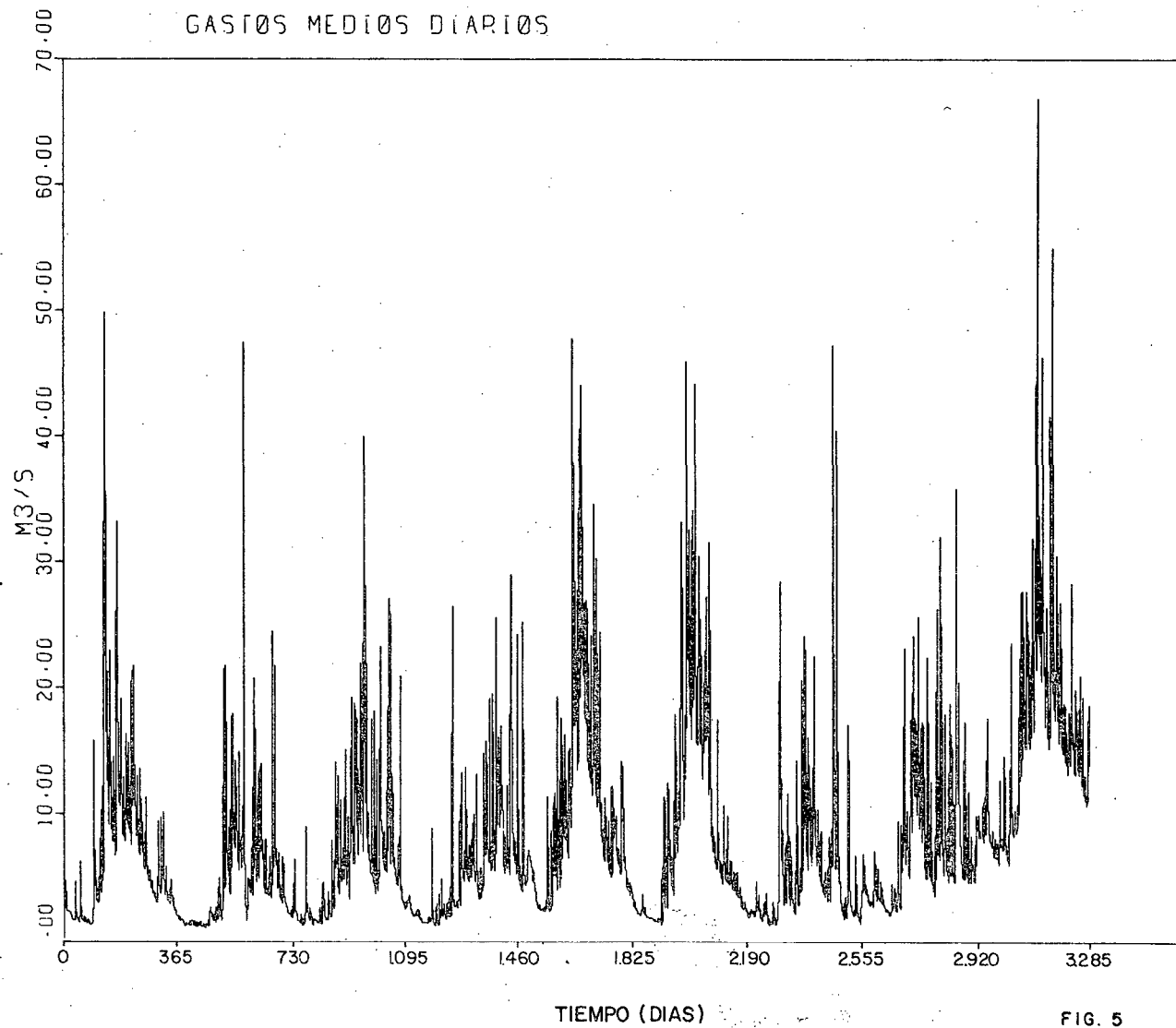


FIG. 5

RIO GUACHE - EN PUENTE VIEJO 1953 - 1966
SERIE DE FOURIER AJUSTADA A LOS VALORES MEDIOS DIARIOS

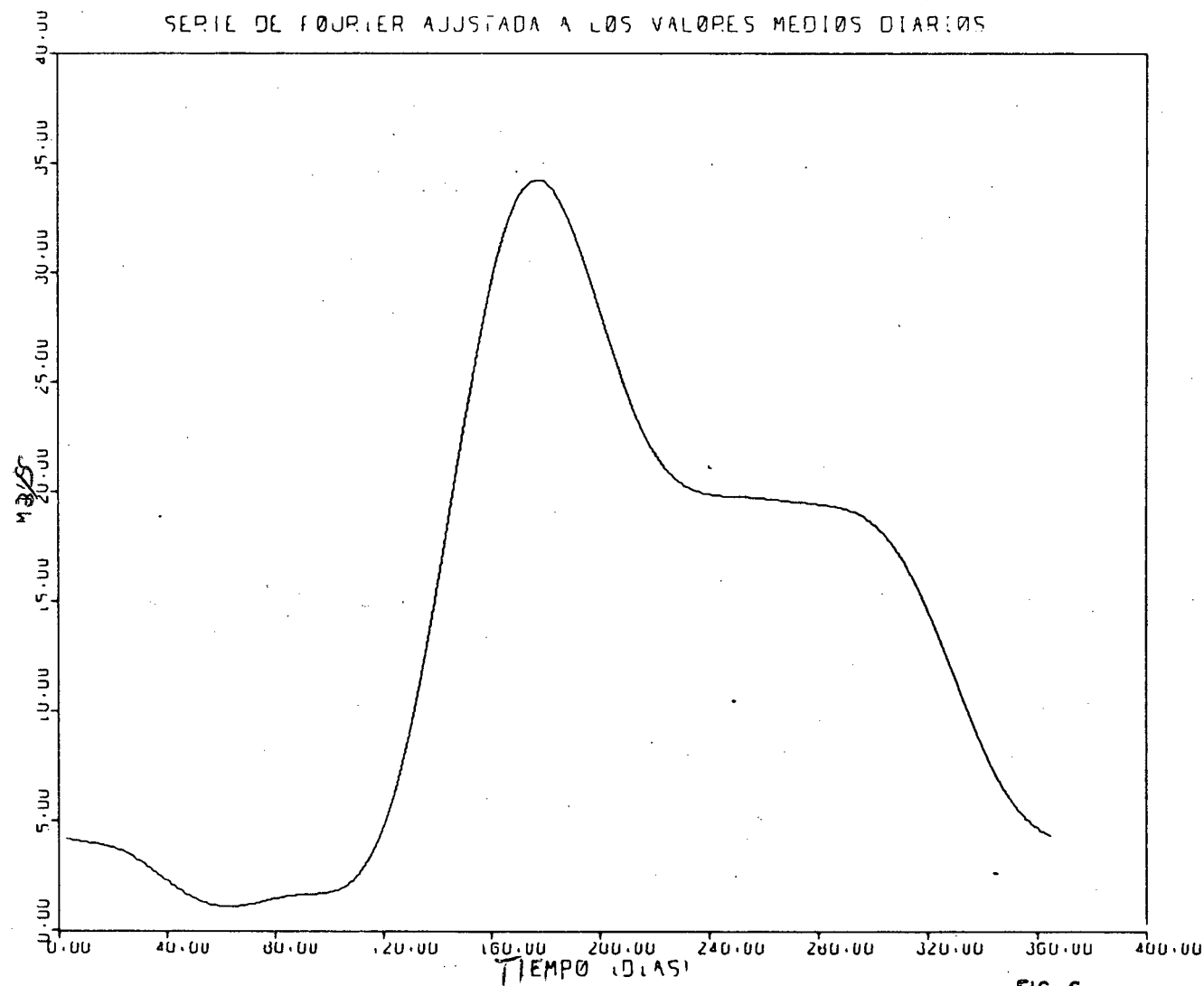


FIG. 6

RIO MOTATÁN - EN AGUA VIVA 1943 - 1969
SERIE DE FOURIER AJUSTADA A LOS VALORES MEDIOS DIARIOS

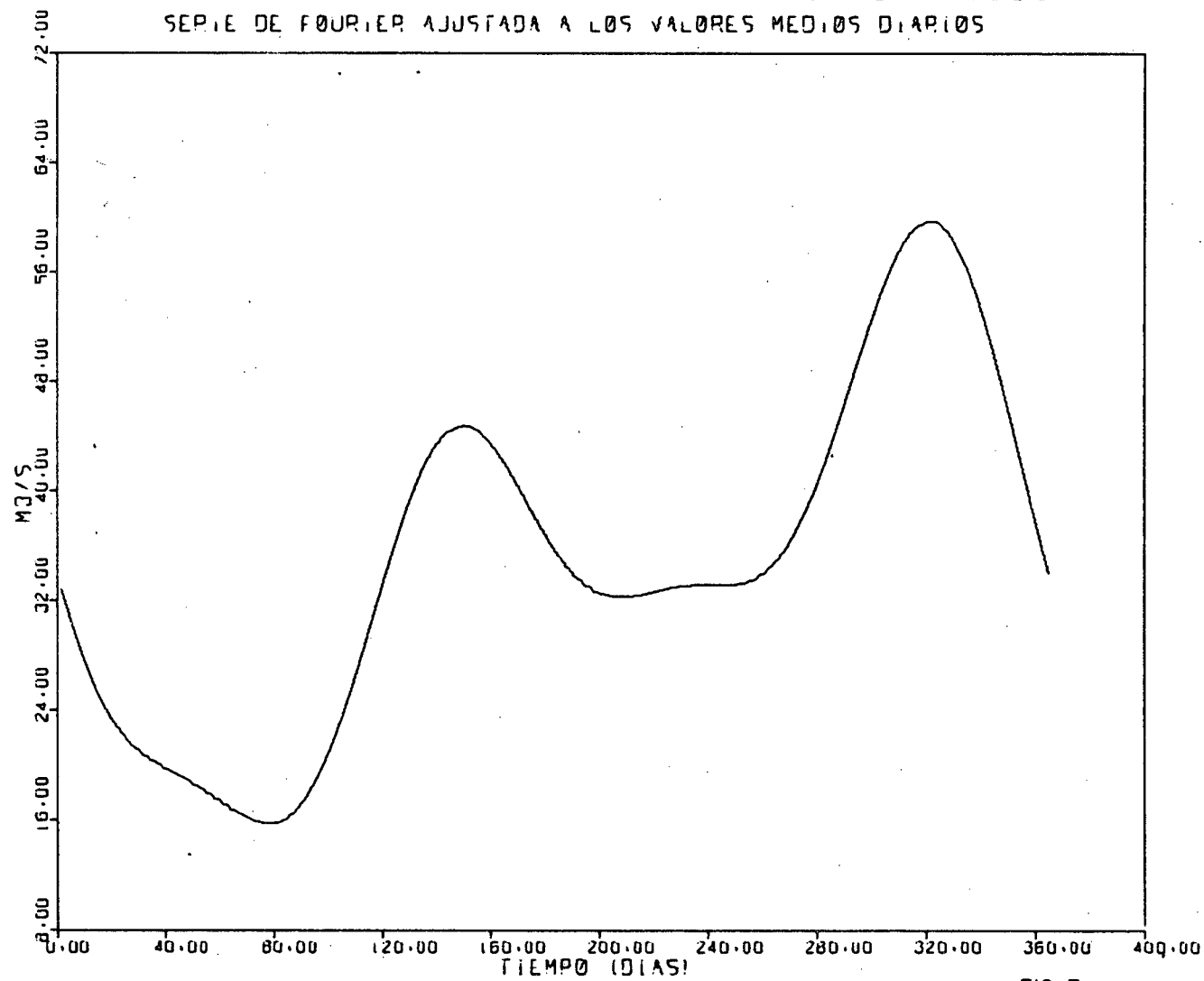


FIG. 7

RI0 PAGUEY - EN EL PASO 1954 - 1966
SERIE DE FOURIER AJUSTADA A LOS VALORES MEDIOS DIARIOS

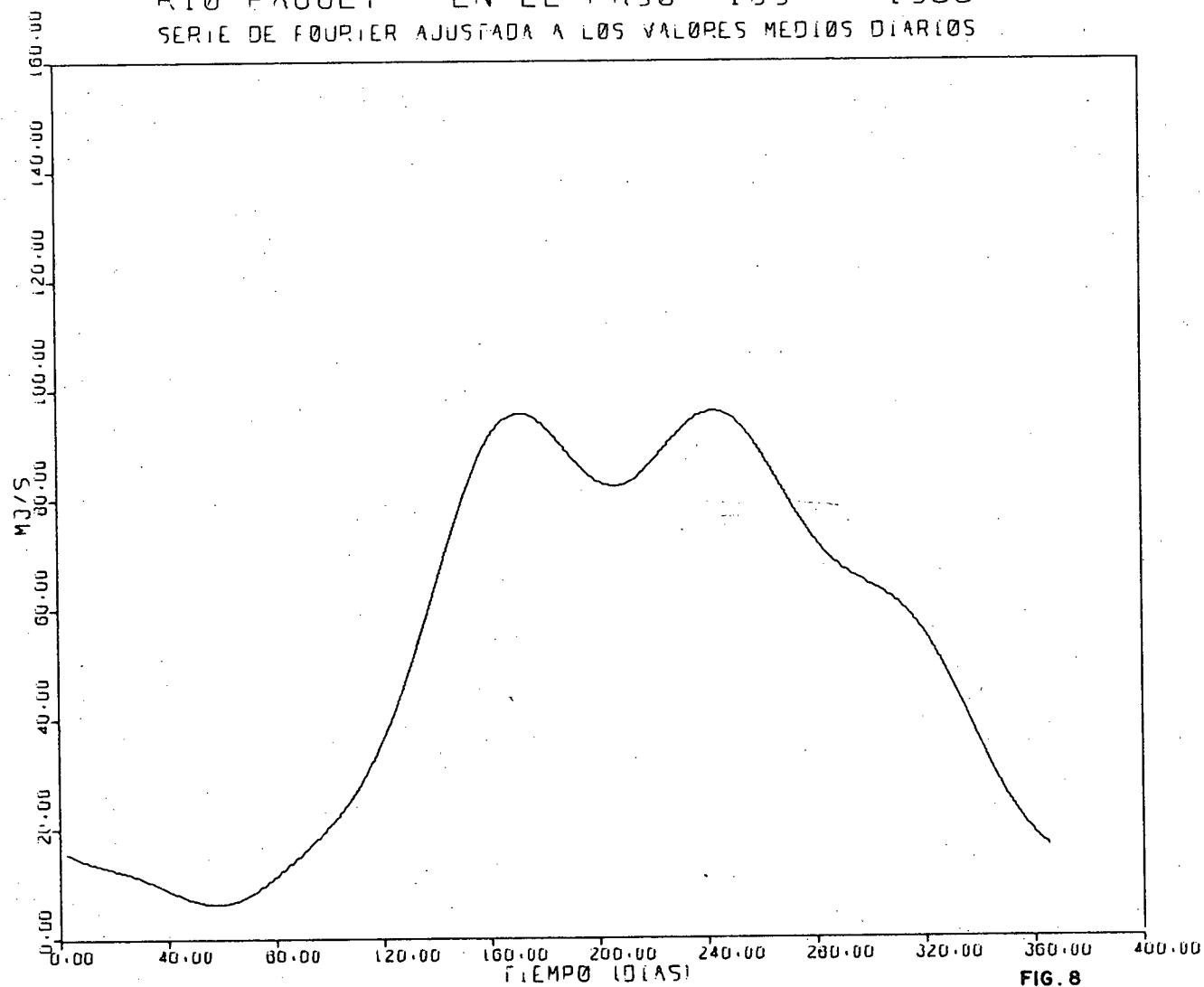


FIG. 8

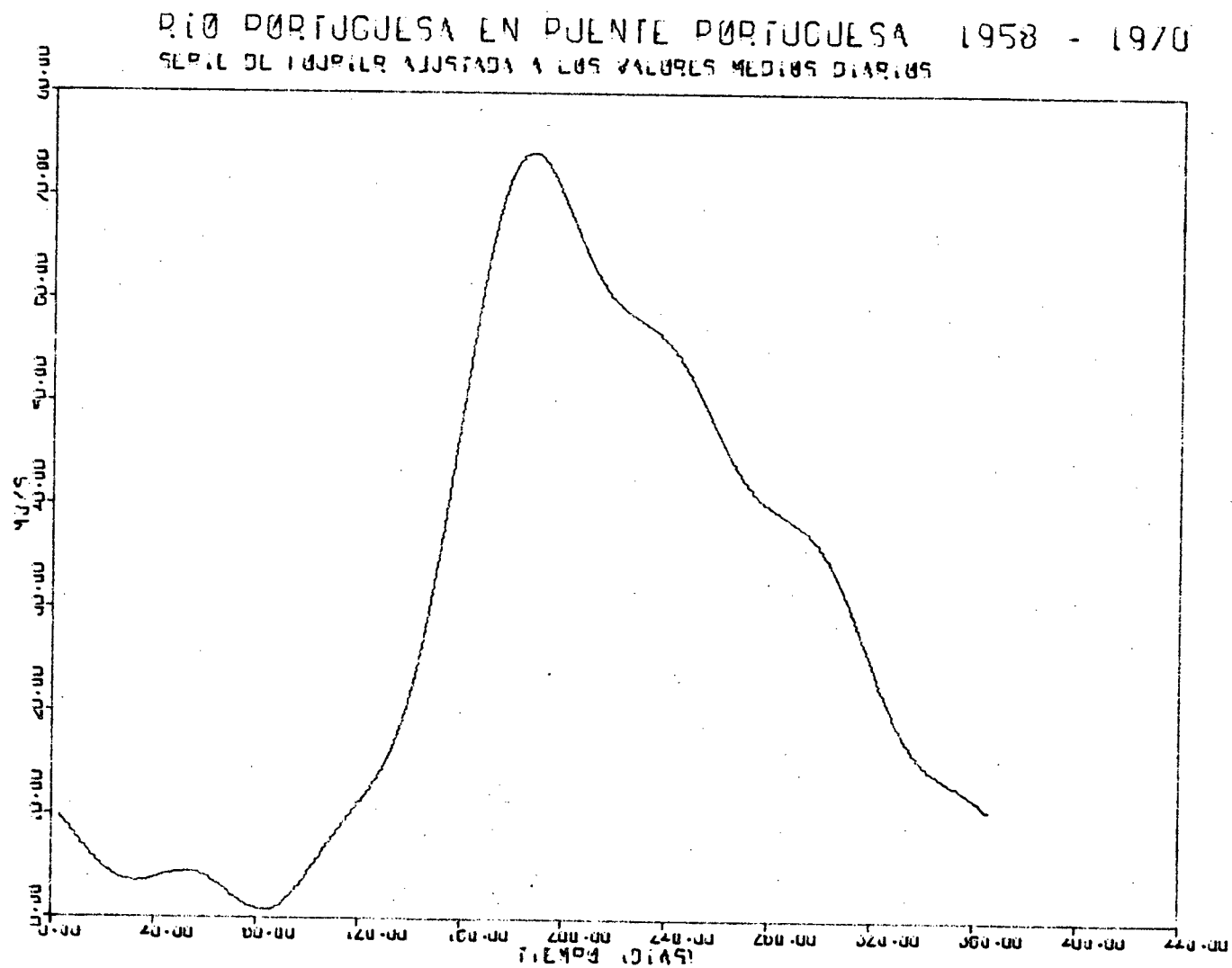


FIG. 9

RIO TORBES - EN SABANETA 1963 - 1971
SERIE DE FOURIER AJUSTADA A LOS VALORES MEDIOS DIARIOS

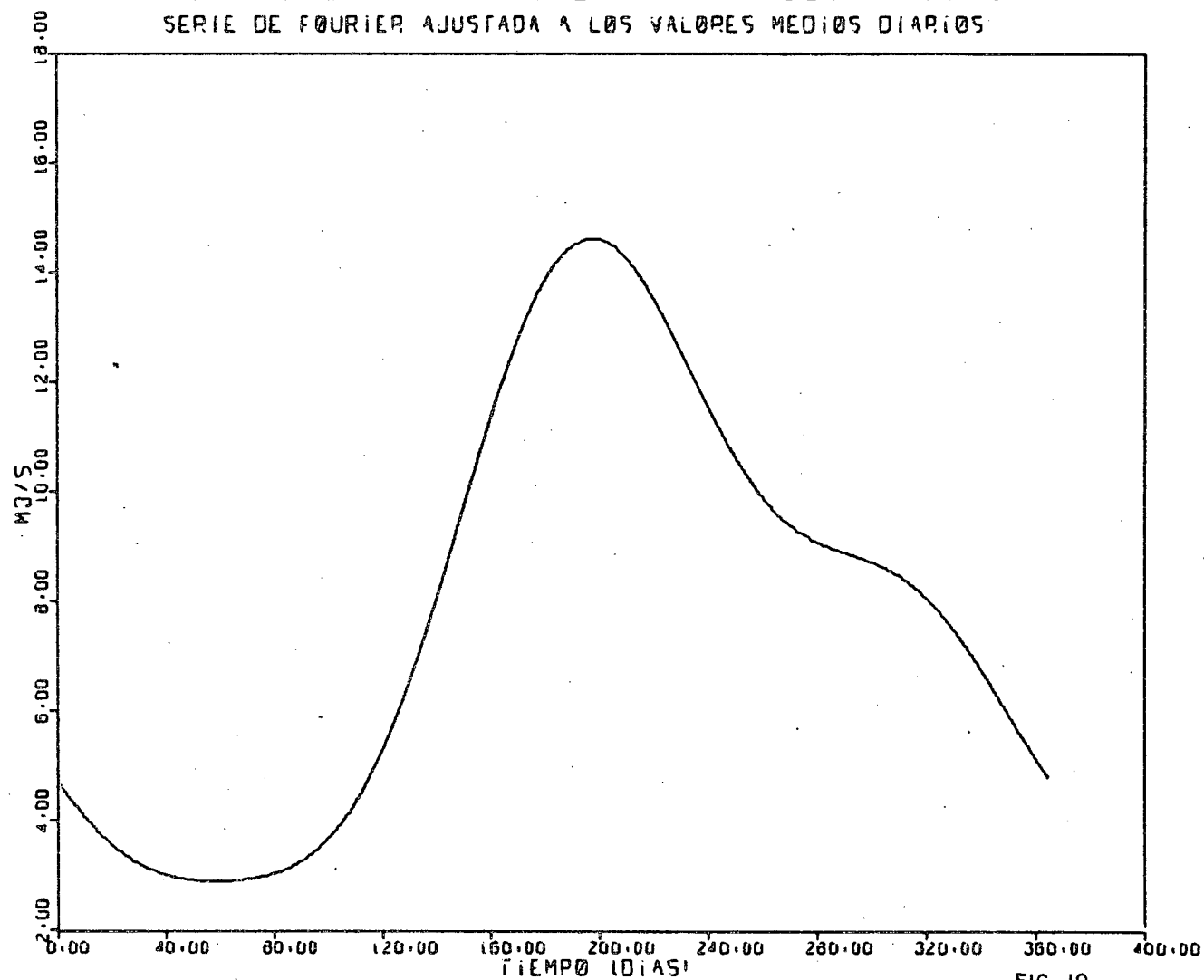


FIG. 10

RIØ GUACHE EN PUENTE VIEJØ 1954 - 1967 CØRRELØGRAMA

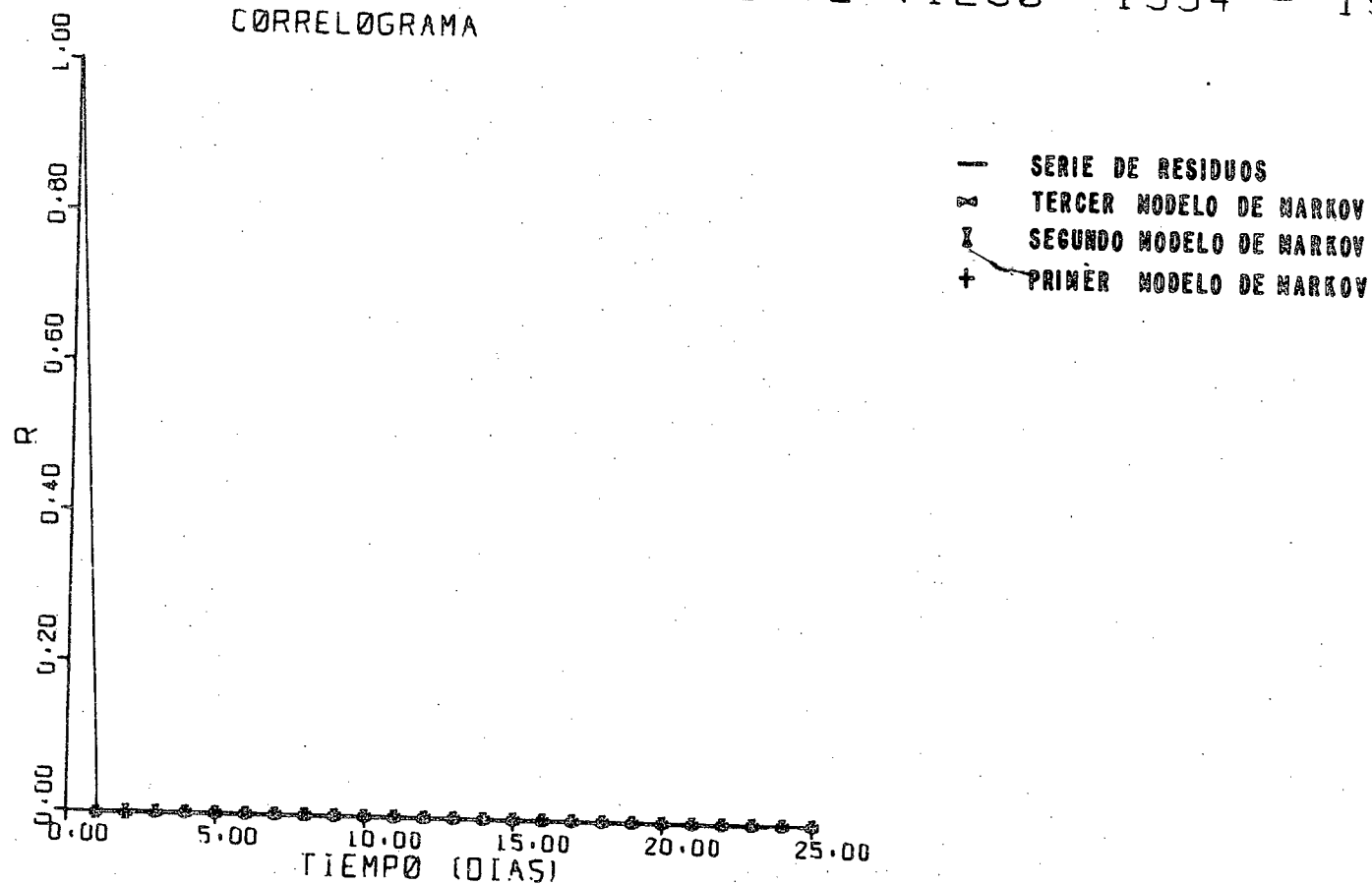


FIG. 11

RIØ MØTATAN EN AGUA VIVA 1943 - 1969 CØRRELOGRAMA

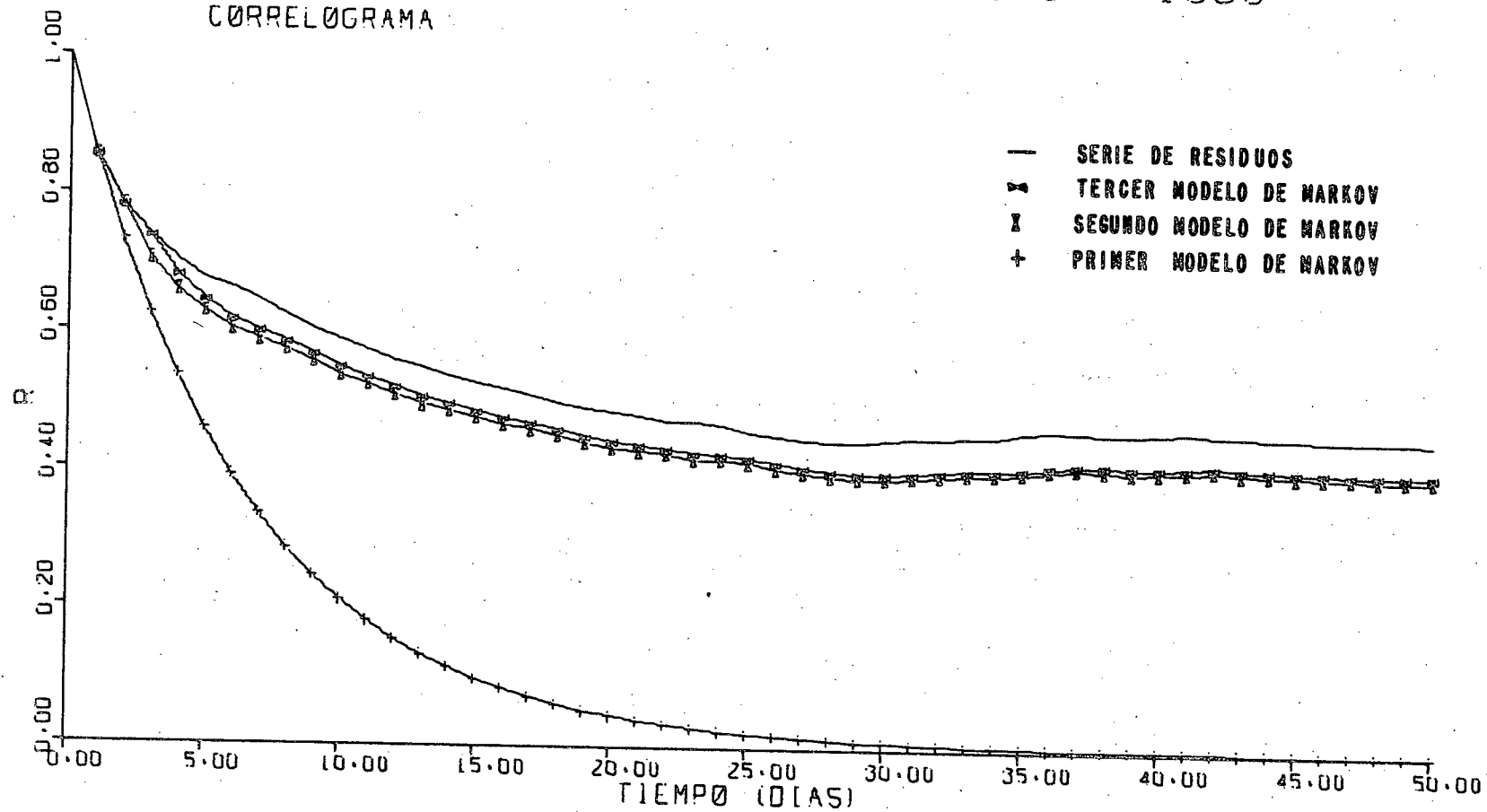


FIG. 12

RIØ PAGUEY EN EL PASØ 1954 - 1966

CØRRELOGRAMA

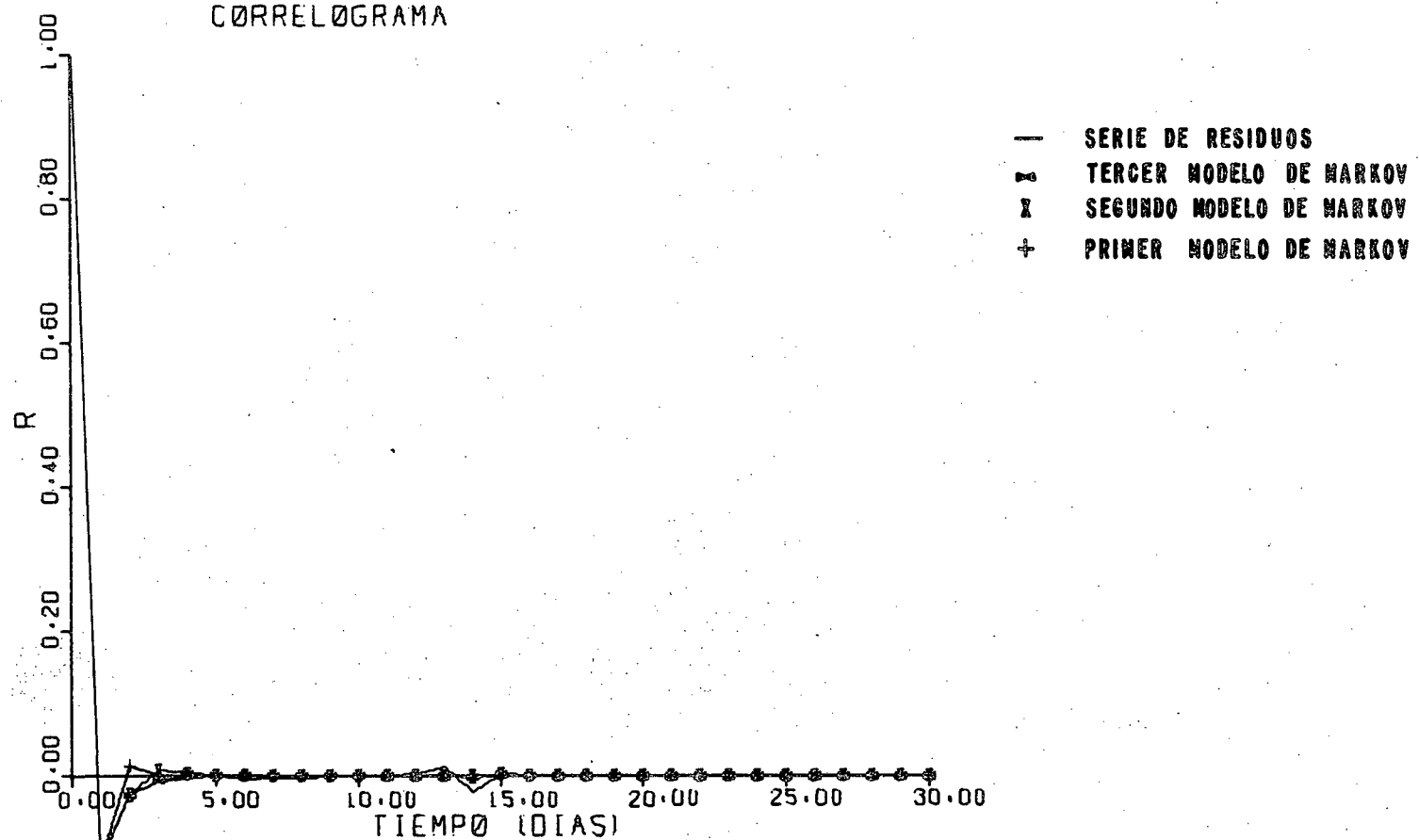


FIG. 13

RIO PORTUGUESA EN PUENTE PORTUGUESA 1958 - 1970

CORRELOGRAMA

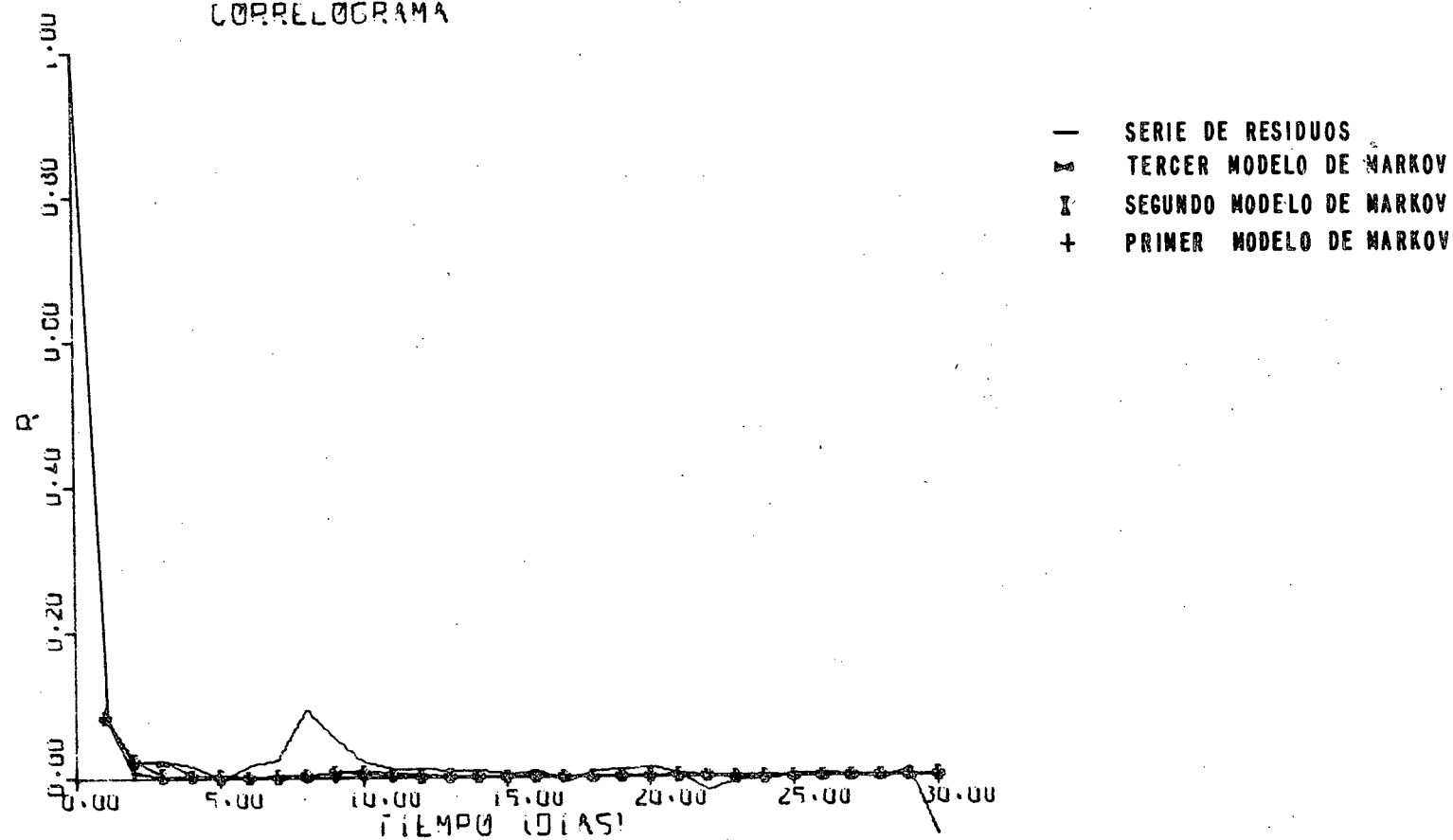


FIG. 14

RIØ TØRBES EN SABANETA 1963 - 1969 CØRRELØGRAMA

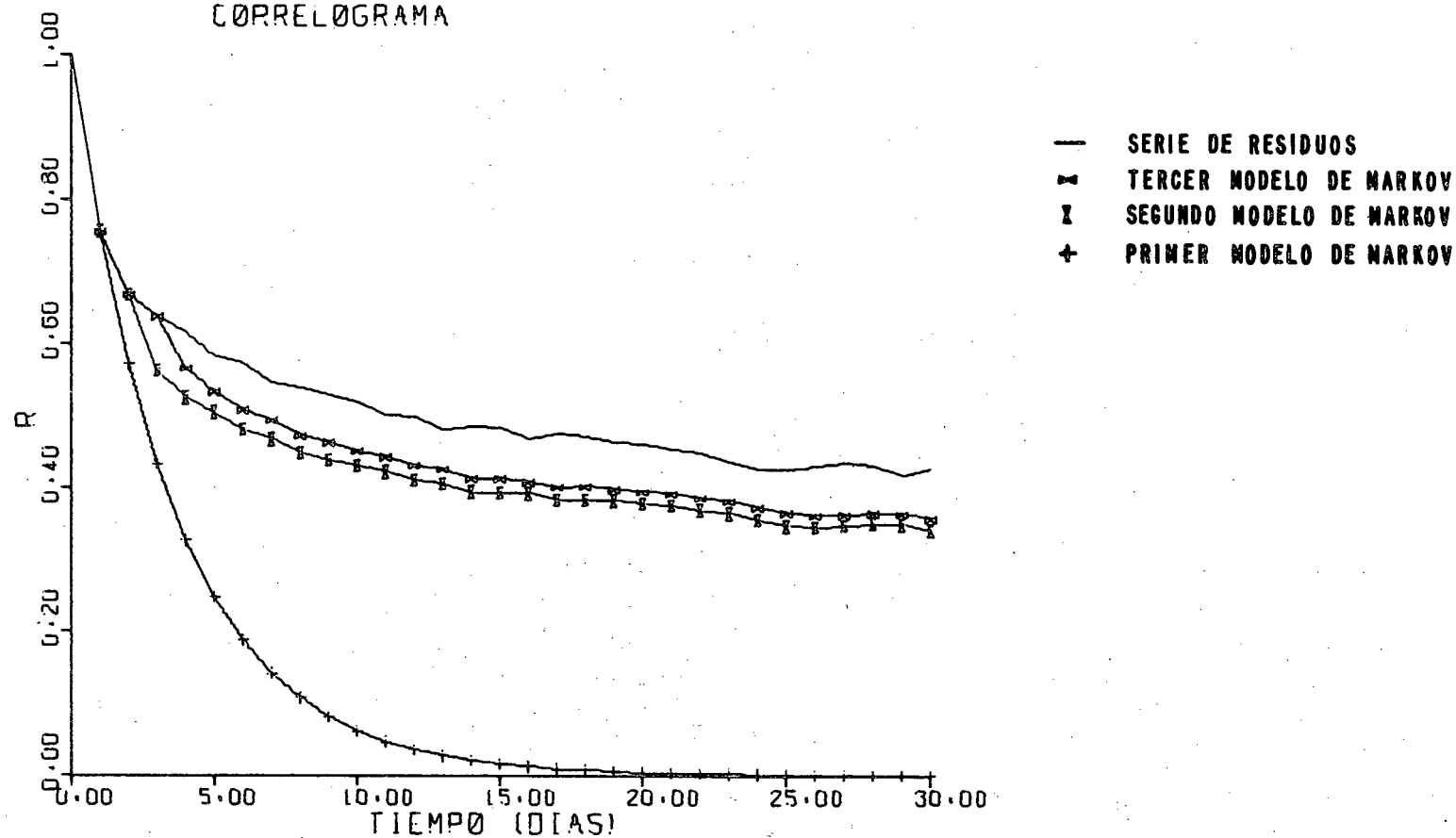


FIG. 15