
SOCIEDAD CHILENA DE INGENIERIA HIDRAULICA

I COLOQUIO NACIONAL

JUNIO-JULIO 1971

SANTIAGO - CHILE

GENERACION ESTOCASTICA DE INFORMACION HIDROLOGICA

Eduardo Varas C. (1)

Este trabajo describe la aplicación de un modelo estocástico de generación de información, de tipo Markoviano en una región, con el objeto de reproducir nuevas series de información hidrológica con iguales propiedades probabilística y estadísticas que las series históricas. Estas series se emplean en estudios de simulación para el diseño de sistemas de recursos de agua. Los resultados obtenidos permiten confirmar el uso de las técnicas de la Hidrología Estocástica en la generación de información.

(1) Profesor Departamento de Obras Hidráulicas, Escuela de Ingeniería, Universidad Católica de Chile.

INTRODUCCION

En los últimos años, la simulación se ha convertido en una importante herramienta de diseño en el análisis de sistemas de recursos de Agua. Este tipo de técnica es muy indicada en estos sistemas, cuya complejidad no permite abordar el problema de maximización en términos generales.

El uso de un modelo de simulación en el estudio de un sistema de recursos de agua, requiere de un registro largo de información hidrológica. Aún más, el estudiar el comportamiento del sistema como una serie de registros hidrológicos, como ser, la serie observada, da sólo una estimación del comportamiento y ello impide formular juicios probabilísticos en relación al sistema. Por ello, existe la necesidad de desarrollar métodos que permitan generar series de información hidrológica tan posibles de ocurrir en el futuro como la serie histórica.

Las variables hidrológicas que condicionan la disponibilidad de agua en el sistema y por consiguiente sus beneficios, están representadas por un conjunto de variables aleatorias cuyos valores en el futuro se desconocen. Este hecho ha obligado al ingeniero a adoptar una de las posiciones siguientes. Al principio, él supuso que la serie histórica disponible se repetiría en el futuro, o que en ella existía un cierto período crítico que podría ser usado en el diseño. De este modo, él podía definir una secuencia de información hidrológica futura. Un enfoque más reciente es aceptar la incertidumbre presente en las variables y suponer que el período histórico define los momentos principales y la ley de probabilidad que ellas siguen. Aceptando estas hipótesis, puede generarse una serie del largo deseado, que tendría las mismas propiedades que la serie histórica y que podría constituirse en una posible serie futura.

Ha sido la labor de la Hidrología Estocástica u

Operacional el desarrollar los métodos que permitan generar esta información sintética con el fin de ser usada como variable de entrada de un modelo de simulación.

ESTUDIOS ANTERIORES

El tipo de modelo empleado en el estudio fue propuesto inicialmente por Thomas y Fiering (7). Básicamente supone que la variable en cualquier período es una función lineal de la variable en el período anterior, incluyéndose, además una variable aleatoria en esta relación de modo de reproducir la varianza existente en los datos hidrológicos. Este esquema puede modificarse fácilmente para incluir períodos estacionales o cíclicos en cada período anual.

Posteriormente Fiering (4) propuso un modelo similar, usando el análisis de componentes principales para resolver el problema de generar información en una región que no sólo mantuviera las propiedades estadísticas y probabilísticas en cada punto a lo largo del tiempo, sino que además se mantuviera la correlación espacial existentes entre estaciones.

Otros enfoques al análisis regional han sido propuestos por Beard (2, 3) y recientemente por Young y Pisano (9). Ellos proponen el uso de técnicas de regresión múltiple y expresan la variable en cada estación como una función de varias variables, como ser, valores de la variable en el mismo período en las demás estaciones, valores anteriores en la misma estación, o ambas.

MODELO MATEMATICO DE GENERACION

El modelo matemático usado en esta generación regional ha sido uno basado en el uso del análisis de componentes principales. Los objetivos principales de este tipo

de análisis son la posibilidad de reducir el número de variables del problema, ordenar las variables de acuerdo a su importancia y eliminar la dependencia espacial entre estaciones. El método se basa en definir un nuevo conjunto de variables o componentes como una combinación lineal de las variables originales estandarizadas, de tal manera que el primer componente principal tenga la máxima varianza, el segundo componente principal se elige ortogonal no correlacionado al primero y con la máxima varianza posible y así sucesivamente. Como resultado de este procedimiento se efectúa una transformación de las variables originales tal que el nuevo conjunto de componentes tiene correlación entre sus elementos y estas variables representan la varianza del espacio muestral en orden de magnitud decreciente.

Los componentes principales se definen de acuerdo a la relación siguiente:

$$Z_i = \sum_{j=1}^p u_{ij} X_j \quad j = 1, 2, 3, \dots, p$$

Siendo :

- p = número de dimensiones del espacio (estaciones)
- u_{ij} = componentes de los vectores propios (elemento j del vector propio i)
- V_i = valores propios de la matriz de correlación espacial
- X_j = variable estandarizada j .
- Z_i = componente principal i con varianza V_i

Este procedimiento permite una reducción en las dimensiones del problema al identificar aquellas variables que son linealmente dependientes del resto o al decidir representar sólo una parte de la varianza de la muestra y trabajar con un modelo de menor número de dimensiones.

Una vez que las variables originales han sido trans

formadas a componentes principales, puede abordarse el problema de generar componentes principales mediante el uso de modelos secuenciales de una variable. Hecho esto, los componentes generados pueden a su vez ser transformados a variables hidrológicas usando las ecuaciones lineales de regresión para cada estación en función de los componentes principales.

Los componentes principales fueron generados usando el siguiente modelo:

$$z_{k,i+1} = \bar{z}_{k,j+1} + B_{k,j} (z_{k,i} - \bar{z}_{k,j}) + S_{k,j+1} (1-R_{k,j}^2)^{\frac{1}{2}} t$$

siendo :

- $z_{k,i+1}$ = valor del componente principal k en el período i+1
- $\bar{z}_{k,j+1}$ = valor medio del componente principal en el mes j+1
- $S_{k,j+1}$ = desviación standard del componente k en el mes j+1
- $B_{k,j}$ = coeficiente de regresión para estimar el valor del componente k en el mes j+1 a partir del valor en el mes j.
- $R_{k,j}$ = coeficiente de correlación temporal entre los valores en el mes j+1 y j del componente.
- t = variable aleatoria standard cuya distribución de las variables generadas.

En el modelo empleado se usó una distribución Gamma, de modo que t tenía la misma distribución y sus parámetros se determinaron poniendo la condición que los tres primeros momentos de la distribución de los componentes principales fueron iguales a los valores históricos.

Finalmente las variables hidrológicas se calcularon

a partir de los componentes generados usando la siguiente relación:

$$X_{k,i} = C_{k1} Z_{1i} + C_{k2} Z_{2i} \dots\dots\dots C_{kn} Z_{ni}$$

en la cual

- X_{ki} = variable en la estación k en el período i
- C_{ij} = coeficiente de regresión entre la variable en la estación i y el componente principal j
- Z_{ki} = valor del componente principal k en el período i.

Las hipótesis del modelo usado son las suposiciones en relación a la distribución probabilística de las variables, las relaciones lineales existentes entre los componentes principales y las estaciones y el modelo de Markov usado para generar componentes principales. Se estima que en el presente caso estas hipótesis son buenas aproximaciones de la realidad.

DIAGRAMA DE FLUJO

El programa de computación desarrollado para llevar a cabo esta generación regional está formada por las etapas que se presentan en el diagrama de flujo de la Figura 1.

Empieza el programa por leer la información hidrológica disponible y calcular las propiedades estadísticas mensuales para cada estación, así como los promedios y desviaciones standard globales en cada estación. Una vez calculados estos estadísticos se procede a estandarizar las variables.

Luego se calcula la matriz de correlación espacial,

es decir, la matriz formada por los coeficientes de correlación lineal entre cada par de estaciones. Esta matriz constituye la información esencial del análisis de componentes principales.

En seguida, se calculan los vectores y valores propios de la matriz de correlación espacial. Teniendo éstos, se puede proceder al cálculo de los componentes principales. Esta matriz de componentes principales, representa toda la información contenida en las variables hidrológicas iniciales.

El paso siguiente, consiste en calcular las propiedades estadísticas de los componentes principales, con el fin de calcular los parámetros necesarios para la generación.

A continuación se establecen los coeficientes de regresión para los valores en cada estación en función de los componentes principales. Estas ecuaciones se usan posteriormente para obtener la información hidrológica a partir de los componentes principales sintéticos.

Sigue la generación de componentes principales, utilizando las relaciones secuenciales descritas en la sección anterior. Dado que los componentes no están correlacionados entre sí, esta generación puede hacerse en forma independiente para cada uno.

Finalmente, los componentes principales generados se transforman en información hidrológica mediante las relaciones lineales. Viene luego el cálculo de las propiedades estadísticas de las variables generadas, las cuales resumen la información sintética y permiten verificar todo el procedimiento.

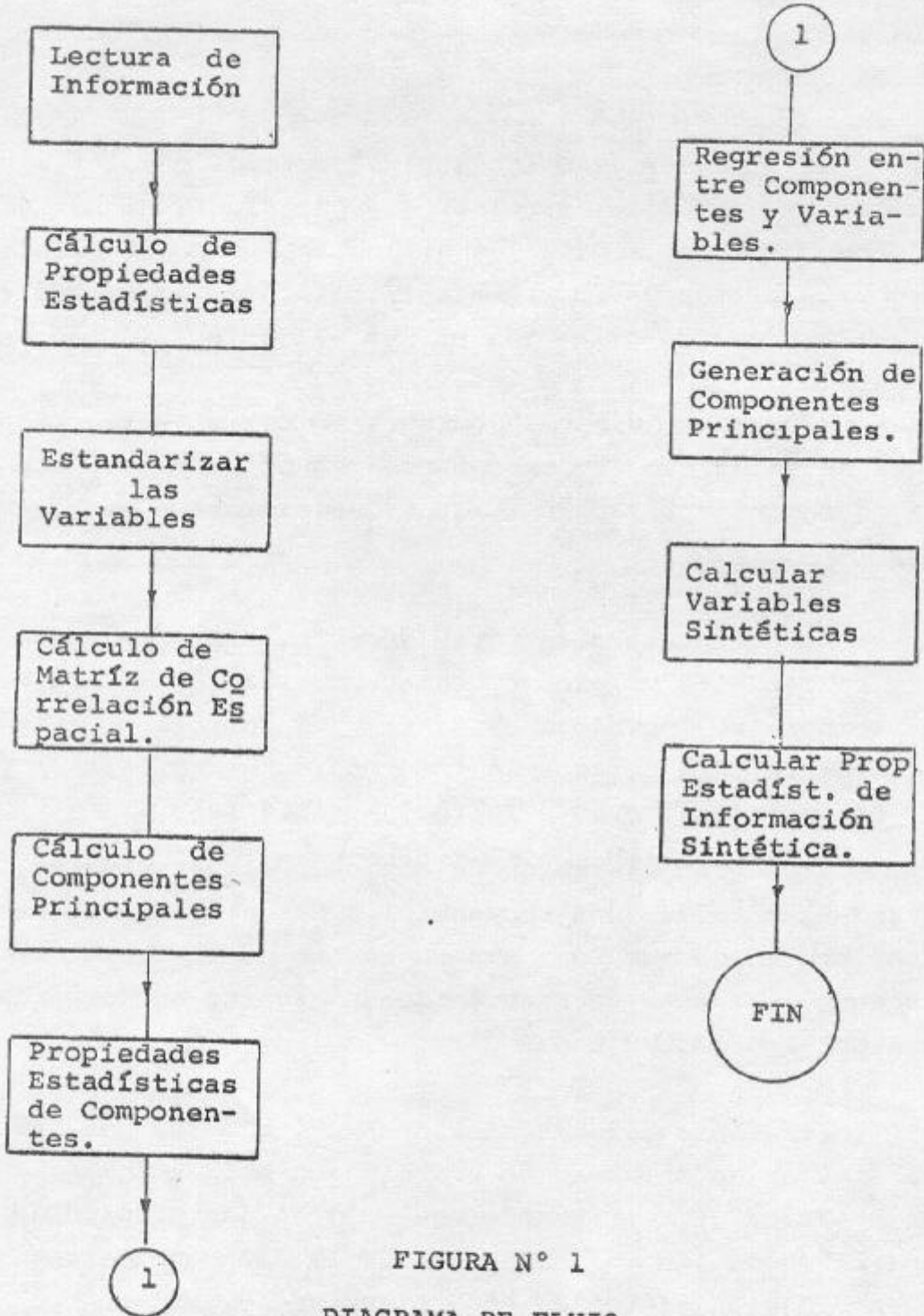


FIGURA N° 1
DIAGRAMA DE FLUJO

ZONA DE ESTUDIO

El modelo antes descrito se aplicó a la Hoya del Río Imperial, ubicada entre los paralelos 38 y 39 de latitud Sur, con una superficie de 12.000 Km².

Esta zona presenta un clima lluvioso, con una precipitación anual de aproximadamente 1300 mm. repartida a lo largo de todo el año. El régimen del río es de tipo pluvial, con una producción específica del orden de 90 l/s/Km².

La Figura N°2 presenta un mapa de la hoya, en el cual se indican las estaciones en las cuales se ha generado información, tanto de precipitaciones totales mensuales como de caudales medios mensuales.

RESULTADOS Y CONCLUSIONES

Se ha encontrado en un estudio anterior, que las variables sintéticas son consistentes con la muestra histórica tanto en la reproducción de los momentos principales de la distribución como también en su forma (8). Análoga conclusión se desprende del presente caso; pudiendo observarse en las tablas que se incluyen la bondad del modelo en la reproducción de los caudales medios y lluvias totales mensuales en las estaciones elegidas.

En términos generales, puede concluirse que es posible generar información hidrológica en varios puntos de una región, con éxito mediante un procedimiento como el señalado. El método mantiene los principales momentos y la forma de la distribución probabilística observada en la serie histórica. Sin embargo, debe tenerse presente que la generación estocástica de información se basa en reproducir probabilística histórica y por consiguiente no añade nueva información, sino que extiende estas probabilidades a nuevas secuencias. Si la muestra histórica no permite estimar los verdaderos parámetros de su distribución, la serie estocástica adolecerá de igual defecto.

BIBLIOGRAFIA

- 1.- ANDERSON, T.W. "An Introduction to Multivariate Statistical Analysis". Wiley, New York (1958)
- 2.- BEARD, L.R. "Use of Interrelated Records to Simulate Streamflows". Jour Hydraulics Div. Am. Soc. of Civil Eng., Vol. 91, pp. 13-22 (Sept. 1965)
- 3.- BEARD, L.R. "Hydrologic Simulation in Water Yield Analysis"., Jour Irrigation and Drainage Div. Am. Soc. of Civil Eng. Vol. 93, pp. 33-42, (Mar. 1967)
- 4.- FIERING, M.B., "A Multivariate Model for Synthesis of Streamflows" Jour Hydraulics Div. Am. Soc. of Civil Eng. Vol. 90, pp. 43-60 (Sept 1964)
- 5.- HUFSCHMIDT, M. y M.B. FIERING. "Simulation Techniques in the Design of Water Resource Systems" Harvard Univ. Press, Cambridge, Mass (1966)
- 6.- KENDALL, M.G. "A Course in Multivariate Analysis". Hafner Publ. Co., New York, (1957)
- 7.- MAASS, A y otros, "Design of Water Resource Systems" Harvard Univ. Press, Cambridge, Mass (1962)
- 8.- VARAS, E. "Sequential Generation of Streamflow" Tesis Ingenier, Univ. de Stanford, California (1968).
- 9.- YOUNG, G.K y W.C PISANO, "Operational Hydrology Using Residuals, Jour Hydraulics Div. Am. Soc. of Civil Eng. Vol. 94, pp. 909-923, (Julio 1968).

GASTOS MEDIOS MENSUALES (M³/s)

HOYA DEL RIO IMPERIAL

	Lumaco en Lumaco (2)		Cautín en Rariruca (9)		Muco en Pte. Muco (10)		Quepe en Vilcún (12)	
	Hist.	Gen.	Hist.	Gen.	Hist.	Gen.	Hist.	Gen.
E	4.8	4.8	68.5	67	9	8	21	20
F	4.1	3	53.2	52	7	5	18	17
M	4.2	4	52.8	54	7	6	18	17
A	5.5	6	57.7	61	10	10	21	22
M	16.1	16	127	127	31	29	45	42
J	39.8	40	165	165	51	53	54	54
J	47.6	48	173	173	56	57	58	57
A	41.9	39	161	154	52	48	54	51
S	35.3	33	147	141	46	43	48	45
O	18.9	18	125	121	28	26	38	37
N	10.0	9	106	103	17	16	30	28
D	6.4	6	90	86	13	12	26	25
Anual	19.5	19	111	109	27	26	36	35

TABLA N° 1.

LLUVIAS TOTALES MENSUALES (mm)

HOYA DEL RIO IMPERIAL

	Victoria		Temuco		Puerto Saavedra		Boroa	
	Hist.	Gen.	Hist.	Gen.	Hist.	Gen.	Hist.	Gen.
E	43	51	40	40	39	40	36	41
F	43	38	40	35	33	31	34	34
M	67	73	64	70	56	62	63	62
A	106	97	104	96	85	81	100	86
M	244	223	199	179	172	161	181	169
J	275	275	200	205	195	197	192	205
J	243	259	180	197	175	180	181	177
A	223	218	158	170	141	138	148	144
S	137	143	100	102	88	93	94	90
O	88	87	76	74	57	59	68	61
N	80	78	66	70	56	60	60	64
D	54	48	56	46	43	39	51	42
Annual	1601	1593	1284	1284	1140	1140	1208	1175.

TABLA N° 2.

