

Com :
V Congreso Nacional de Hidráulica
Asociación Ecuatoriana de Hidráulica,
AEHID/CICP, Quito, 23-26/11/1988.

HOMOGENEIZACION Y REGIONALIZACION PLUVIOMETRICA POR EL METODO DEL VECTOR REGIONAL

por Patrick Le Goulven*, Miguel A. Alemán**, Iván Osorno**

RESUMEN

Después de haber definido el concepto de homogeneidad de las series de precipitaciones anuales y demostrado la ineficacia de las pruebas estadísticas clásicas para resolver este problema, los autores demuestran la eficiencia de las pruebas de dobles masas y sobre todo del vector regional.

Utilizan además este último método para emprender una regionalización pluviométrica teniendo en cuenta la variación interanual de las precipitaciones (regionalización cronológica) y su repartición a lo largo del año (regionalización de régimen).

Las demostraciones y conclusiones se ilustran en base a trabajos ya realizados en la cuenca hidrográfica del río Mira.

* Ingeniero hidrólogo del Instituto Francés de Investigación Científica para el Desarrollo en Cooperación (ORSTOM)

** Ingenieros hidráulicos del Instituto Ecuatoriano de Recursos Hidráulicos (INERHI)

1. The first part of the document discusses the importance of maintaining accurate records of all transactions.

2. It is essential to ensure that all entries are supported by appropriate documentation and receipts.

3. Regular audits should be conducted to verify the accuracy of the records and identify any discrepancies.

4. The second part of the document outlines the procedures for handling and storing financial records.

5. Records should be organized in a systematic manner to facilitate easy access and retrieval.

6. Adequate security measures should be implemented to protect the confidentiality and integrity of the data.

7. The final section provides a summary of the key points and offers recommendations for further improvement.

8. It is recommended that the organization review these guidelines periodically to ensure they remain relevant and effective.

9. The document concludes by emphasizing the commitment to transparency and accountability in all financial operations.

10. For more information, please contact the Finance Department at [contact information].

INTRODUCCION

El inventario de los recursos hídricos así como su manejo racional son preocupaciones actuales del Ecuador.

Sin embargo, tales objetivos dependen estrechamente de un buen conocimiento de los elementos del clima entre los cuales la pluviometría juega un papel importante. Su conocimiento cuantitativo se obtiene por mediciones permanentes a través de una red de observaciones puntuales.

Lastimosamente, todo registro de datos adolece de errores de toda naturaleza, lo que demuestra la necesidad de una crítica severa de los datos obtenidos por medio de la red, los mismos que pueden ser eliminados cuando se alejan demasiado de la realidad, corregidos cuando la naturaleza de los errores lo permite, o simplemente señalados cuando el valor anormal puede ser considerado como localmente posible.

Además, el cálculo de las necesidades de agua y de los recursos disponibles requiere una información completa y consistente aunque no existan estaciones cercanas.

A fin de alcanzar este objetivo, es indispensable dividir el espacio estudiado en zonas homogéneas, en las que el régimen pluviométrico así como la variación interanual de la precipitación sean similares y bien conocidos, para después diseñar modelos hidro-pluviométricos que permitan manejar con criterio científico el recurso agua.

El Instituto Ecuatoriano de Recursos Hidráulicos (INERHI), en colaboración con el Instituto Francés de Investigación Científica para el Desarrollo en Cooperación (ORSTOM), emprendió esa tarea en la Sierra y en la Costa, para elaborar posteriormente su Plan Nacional de Riego. Ante la magnitud del trabajo, se asociaron al programa la Escuela Politécnica Nacional (EPN) y el Instituto Nacional de Meteorología e Hidrología (INAMHI).

Las cuatro instituciones trabajan en regiones diferentes en base a una metodología común que se presenta en este artículo.

A - HOMOGENEIZACION

1. DEFINICION DEL PROBLEMA

Una serie cronológica de valores observados es de carácter aleatorio sencillo si todos sus valores son independientes y provienen de una misma población estadística por sorteo al azar. El carácter aleatorio puede ser alterado por :

- **un efecto de persistencia**
Los valores son dependientes de los anteriores pero la serie es estacionaria y sus parámetros de distribución no varían en el tiempo.
- **un efecto de tendencia**
La esperanza matemática de los valores observados varía en el tiempo.
- **efectos cíclicos o pseudo-cíclicos**
La esperanza matemática de un valor observado es función de la cronología, pero el valor promedio de series suficientemente largas (medias móviles) es estacionario y sus parámetros no varían en el tiempo.
- **errores de observación y de medición**
Afectan uno o varios términos consecutivos de la serie y pueden superponerse a los 3 efectos anteriores.

Una serie cronológica de precipitaciones anuales puede contener todos los tipos de efectos al mismo tiempo. Los tres primeros caracterizan a la pluviometría del sitio de medición y conforman así la organización interna de la serie.

Sólo los errores de observación y de medición (o errores sistemáticos) inciden en la homogeneidad de la serie.

El verdadero problema de la homogeneización consiste entonces en detectar e identificar dichos errores.

Tal es el objetivo de esta primera parte.

2. LAS FUENTES DE HETEROGENEIDAD Y LAS CONSECUENCIAS DE ELLA

Múltiples son los problemas que se presentan en una red de observación pluviométrica. Recordaremos solamente los más frecuentes:

- **Modificación del medio circundante o del sitio de medición**
 - Por desplazamiento del aparato
Es un caso frecuente y generalmente la estación conserva el mismo nombre.
 - Por modificación del medio circundante
Puede ser brusca (construcción cercana) o progresiva (crecimiento de árboles a proximidad).
 - Por cambio de la altura del aparato
Modifica significativamente la cantidad de agua recibida por el pluviómetro sobre todo cuando el sitio está expuesto al viento.

- **Errores debidos al aparato**

- modificación de la superficie receptora por construcción o deformación del aparato de medición;
- defecto de calibración (en los pluviógrafos).

- **Errores de medición**

Los errores a nivel de la probeta o de la regleta son frecuentes. La confusión de probeta es el caso típico. Antiguamente, el Sur del Ecuador estaba equipado con pluviómetros de 100 cm² con la correspondiente probeta de cristal. Una vez rota, el observador podía recibir una nueva de plástico, pero calibrada para un pluviómetro de 200 cm². Las mediciones posteriores al cambio subestimaban entonces en un 50 % la altura real de lluvia.

El siguiente cuadro presenta el valor de los errores cometidos en el caso de una confusión entre probetas y pluviómetros de superficies usuales.

		probeta			
		100 cm ²	200 cm ²	314 cm ²	400 cm ²
P l u v i ó m e t r o	100 cm ²	1	0,5	0,318	0,25
	200 cm ²	2,0	1	0,637	0,50
	314 cm ²	3,14	1,57	1	0,79
	400 cm ²	4,0	2,0	1,27	1

Cuadro nº 1

- **Errores de transmisión**

Problemas de correo, de radio, de satélite, etc.

- **Errores de archivo y publicación**

Al tipear la información en bruto, se generan errores de todo tipo (omisión de datos, errores de fecha, de copia, de digitación, etc.).

Según su naturaleza, los errores sistemáticos tienen consecuencias diferentes tanto en su detección como en la corrección de los valores erróneos.

Una confusión de probeta necesita generalmente pocos años para detectarse cuando no hay utilización concomitante de varias probetas o regletas. Es posible en ese caso restablecer los totales anuales, mensuales y hasta diarios cuando se conoce la fecha de la confusión.

Una modificación del medio circundante se detecta con un buen número de años homogéneos antes y después del cambio. Mientras mayor es el cambio, menos años se requieren. La experiencia demuestra que solamente los totales anuales y mensuales pueden corregirse en este caso.

Si hay desplazamientos frecuentes del aparato (u otras causas que producen el mismo efecto), las heterogeneidades no son discernibles y es imposible efectuar correcciones.

3. IMPORTANCIA DE LAS PRUEBAS DE HOMOGENEIDAD

Todo lo anterior demuestra la necesidad de controlar rigurosamente los datos pluviométricos, control que se podría hacer consultando el historial de las estaciones.

En general, un cambio de sitio coincide con un cambio de observador y una confusión de probetas se produce después de una interrupción de las mediciones. Un buen historial con anotaciones sobre cambios de observadores o de sitio, de material o del medio circundante, permitiría entonces resolver gran número de problemas.

Lastimosamente, si bien los inspectores conocen esos detalles de la vida de una estación, estas figuran rara vez por escrito en algún documento, y con la desaparición de « la gente que sabe » desaparece también la información.

Una parte del historial puede reconstruirse en base a los documentos originales enviados por los observadores, pero ello es generalmente insuficiente.

Por tal razón, surgió un interés general por elaborar pruebas de homogeneidad a partir de los simples datos.

Numerosas pruebas estadísticas existen en la literatura: test de los signos, test de correlación serial de Wald, Wolfowitz y Anderson, test de los picos y de las fases, de correlación de rango de Mann y Sperman, de Fisher-Yates, etc.

Esos tests solamente pueden comprobar el carácter aleatorio de una serie de variables independientes, lo cual implica la hipótesis de la inexistencia de una organización interna de la muestra observada.

Por construcción, comparan una parte de los valores de la muestra con la restante, por lo que requieren conocer de antemano la época separando las series homogéneas.

Para resumir, no pueden distinguir los errores sistemáticos de una organización interna real de la serie, pues no han sido concebido para tal efecto.

Desafortunadamente, la mayoría de series pluviométricas anuales tiene una organización interna, lo que disminuye aún más el interés de aplicar a la pluviometría las pruebas clásicas de homogeneidad.

4. LAS PRUEBAS DE SIMPLES MASAS

Para utilizar este método, se anota en coordenadas:

- el número i del año en la serie (creciente o decreciente), omitiendo los años faltantes;
- la suma $\sum x_i/x_m$ en la cual x_i representa las precipitaciones anuales y x_m la pluviometría promedio de todos los años utilizados.

En el caso de una serie cronológica de carácter aleatorio sencillo, el gráfico de simples masas oscilará alrededor de la primera bisectriz.

En la práctica, siendo el trabajo muy largo, se usan programas que utilizan una variable transformada de tipo $x_i/x_m - i$, lo que permite transformar la primera bisectriz en un eje vertical, facilitando así el trazado de los gráficos por computador.

Si la serie ha sido afectada por uno (o varios) de los efectos enumerados en el punto 1, el gráfico presentará un cambio de pendiente significativo. Dicho cambio es fácilmente detectable y permite además localizar los años afectados, lo que ni siquiera podían realizar los tests clásicos.

Lastimosamente, las pruebas de simples masas no pueden distinguir con certeza los errores sistemáticos de eventuales efectos que componen la organización interna de la serie observada.

La figura 1, elaborada automáticamente a partir del programa MASA utilizado en el proyecto, ilustra bien la afirmación anterior.

El gráfico de simples masas de la estación Mira no oscila alrededor del eje vertical y parece tener 3 periodos de pendientes diferentes.

El período 1965-1968 es casi paralelo al eje lo que significa que esta compuesto por años normales (cercanos al promedio multianual).

El período 1969- 1976 corresponde a una serie de totales anuales elevados y el último período (1977-1983) a una serie de precipitaciones « bajas » (inferiores al promedio multianual).

El caso de la estación Los Cerros (Montecristi), ubicada en la provincia del Manabí, es aún mas espectacular.

El gráfico está compuesto de 3 períodos principales (1964-1971, 1972-1976, 1977-1982) y se observa claramente el valor anormalmente elevado del año 1983 que altera completamente el gráfico y que representa evidentemente al fenómeno del Niño.

Todos esos cambios de pendiente corresponden a efectos que vienen a alterar el carácter aleatorio de las dos series, pero el problema reside en saber cuáles son esos efectos.

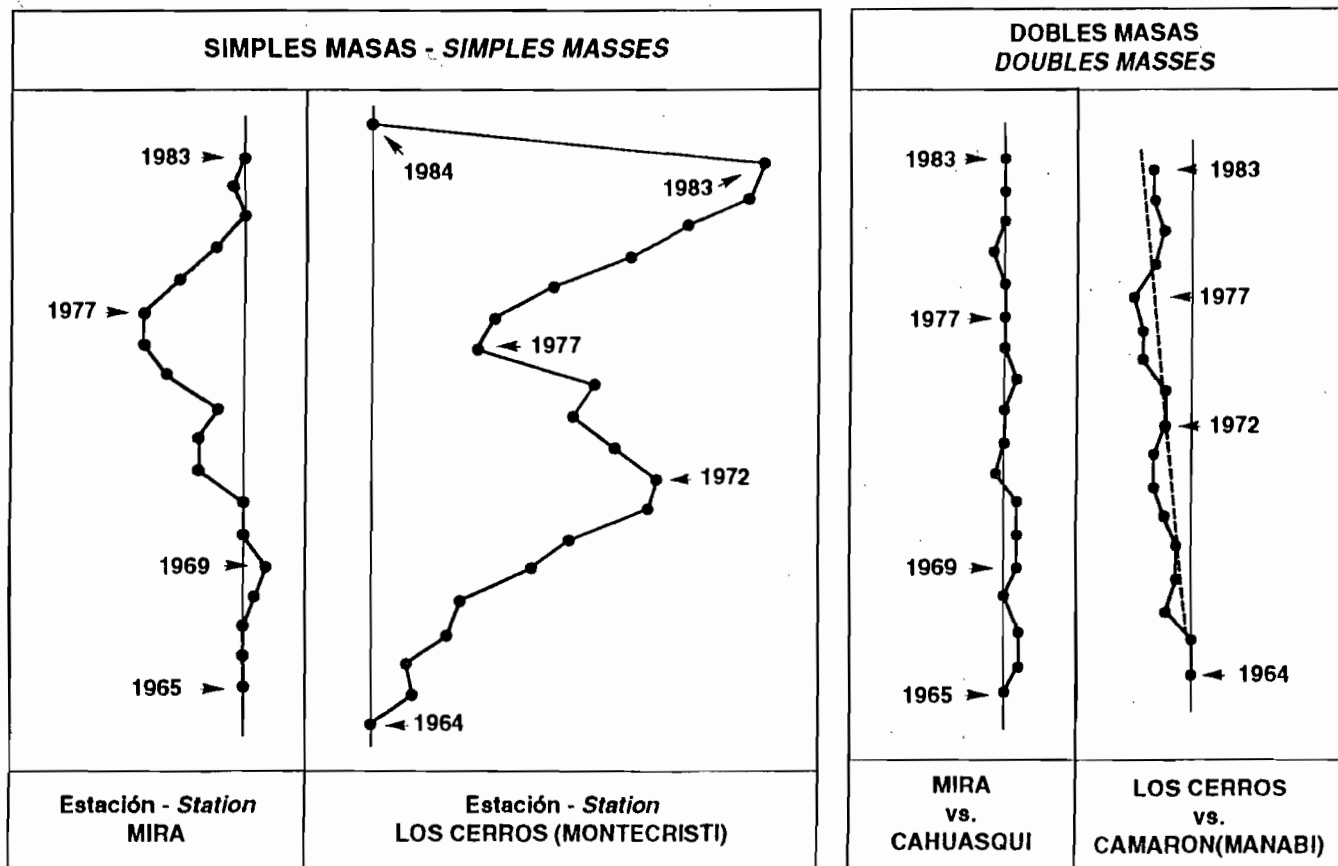


Figura 1

Figura 2

El solo uso de las simples masas no permite contestar objetivamente y la única observación que puede hacerse es que en las 2 estaciones existe un cambio similar de pendiente entre los años 1976 y 1977. Este cambio debe corresponder entonces a un fenómeno normal de organización interna.

Para concluir, es indispensable consultar los gráficos de dobles masas (figura n 2), realizadas con estaciones vecinas y en los cuales han desaparecido los diferentes períodos, lo que demuestra que todos los cambios de pendiente anteriores son debidos a fenómenos normales y característicos de las regiones climáticas en las cuales están ubicadas las 2 estaciones.

Si bien las pruebas de simples masas son interesantes en comparación con las pruebas estadísticas, no pueden aportar criterios realmente objetivos en cuanto a la homogeneización de los datos pluviométricos.

Además parece imposible detectar con certeza los errores sistemáticos a partir de una sola serie cronológica. Afortunadamente, en el estudio de las precipitaciones, se dispone generalmente de varias estaciones ubicadas en una misma región.

5. LAS PRUEBAS DE DOBLES MASAS

5.1 Hipótesis de base

Se admite que existen zonas climáticas en las cuales se dispone de varias estaciones cuyas series cronológicas están relacionadas por coeficientes de correlación lineal significativamente positivos.

Perteneciendo a una misma zona climática, las estaciones están sometidas a los mismos efectos de persistencia, de tendencia o de pseudo-ciclos. Sus coeficientes de variación son entonces similares y sus precipitaciones anuales pseudo-proporcionales.

Es decir que para un año i , los totales anuales x_i y y_i de dos estaciones de la misma zona tienen entre ellos la siguiente relación:

$$(1) \quad x_i/x_m = y_i/y_m + e_i$$

en donde x_m y y_m corresponden a los promedios multianuales de las dos estaciones, calculados en un período suficientemente largo. La varianza de e_i (término independiente tanto de x_i como de y_i) es más reducida si el coeficiente de correlación entre las dos estaciones es mayor. Su esperanza matemática es igual a cero.

5.2 Teoría de las dobles masas

La hipótesis anterior implica entonces que una estación X está ligada a una tendencia climática regional única, cuyos valores anuales t_i pueden expresarse mediante una variable w_i reducida y normada de la siguiente forma:

$$(2) \quad w_i = (t_i - t_M)/\sigma_t$$

en donde t_M corresponde al promedio de la tendencia y σ_t a su desviación estándar.

Considerando que una estación tiene factores de sitio propios (promedio x_M , desviación estándar σ_x), representativos del sitio de medición, y que sus observaciones tienen fluctuaciones aleatorias (v_i) cuya magnitud varía según la relación existente entre la estación y su tendencia, las precipitaciones x_i de la estación X podrán escribirse según la fórmula:

$$(3) \quad x_i = x_M + \sigma_x \cdot (r_x \cdot w_i + u_i \cdot \sqrt{1-r_x^2})$$

en donde r_x representa el coeficiente de correlación entre la estación y su tendencia.

Acumulando los datos hasta el año n , tenemos:

$$(4) \quad \sum x_i = n \cdot x_M + \sigma_x \cdot (r_x \cdot \sum w_i + \sum u_i \cdot \sqrt{1-r_x^2})$$

Dividiendo para el promedio:

$$(5) \quad X_n = \sum x_i / x_M = n + C_x (r_x \cdot \sum w_i + \sum u_i \cdot \sqrt{1-r_x^2})$$

en donde C_x corresponde al coeficiente de variación de la estación X . El gráfico de simples masas ilustra esta relación.

Consideremos ahora otra estación, Y , que pertenece a la misma zona climática. Efectuando los cálculos anteriores, obtendremos la siguiente ecuación:

$$(6) \quad Y_n = n + C_y \cdot (r_y \cdot \sum w_i + \sum v_i \cdot \sqrt{1-r_y^2})$$

Eliminando n y reagrupando los términos:

$$(7) \quad Y_n = \begin{cases} X_n \\ + \sum w_i \cdot (r_y \cdot C_y - r_x \cdot C_x) \\ + (C_y \cdot \sum v_i \cdot \sqrt{1-r_y^2} - C_x \cdot \sum u_i \cdot \sqrt{1-r_x^2}) \end{cases}$$

La ecuación anterior muestra la existencia, entre los totales acumulados X_n y Y_n , de una relación lineal alterada por 2 términos complementarios, A y B .

Eso significa que el gráfico de dobles masas (o de dobles acumulados) oscilará alrededor de una recta de pendiente positiva igual a 1.

Si las estaciones son independientes, la magnitud de la oscilación del gráfico dependerá únicamente de C_x y C_y . En el caso contrario, la oscilación disminuirá cuando los coeficientes de correlación se aproximen a 1.

Toda variación de uno de los promedios se traducirá en una modificación simultánea de la pendiente del gráfico, como en el caso de las simples masas.

Según la ecuación (7), si las 2 estaciones son suficientemente cercanas como para obedecer estrictamente a una misma tendencia climática, C_x y C_y serán sensiblemente iguales así como r_x y r_y . En este caso, el término A de la ecuación tenderá hacia cero y la relación existente entre las dos sumas acumuladas X_n y Y_n será independiente de eventuales fluctuaciones climáticas y alterada únicamente por factores aleatorios cuando los datos son homogéneos.

EN ESTE CASO, TODA MODIFICACION DE LA PENDIENTE DEL GRAFICO CORRESPONDERA UNICAMENTE A ERRORES SISTEMATICOS.

Es la gran ventaja que tienen las dobles masas en comparación con las simples masas cuya ecuación (5) contiene siempre los valores w_i relativos a la tendencia climática de la zona considerada, los mismos que pueden contener otros efectos superpuestos como lo hemos visto en la figura 1.

Este valor no corresponde a ningún valor del cuadro 1, habiéndose diagnosticado entonces un cambio de sitio.

Multiplicando el primer período por el valor de 1,2 para obtener una serie cronológica homogénea y representativa del último sitio del aparato, obtenemos una curva mucho más adecuada.

Sin embargo, después de haber visto la estación y estudiado su historial en el INAMHI (a quien presentamos nuestro reconocimiento por habernos permitido la consulta), pudimos comprobar el cambio de sitio de la estación climatológica, pero en una fecha anterior.

Este error en el diagnóstico proviene del valor particularmente bajo del coeficiente de correlación (0,56) y viene a corroborar lo dicho en el párrafo anterior.

Las dobles masas entre las estaciones Zuleta y Pimampiro muestran un cambio de pendiente a partir del año 1972. La relación entre las pendientes de los dos períodos es de 1,36, lo que tampoco corresponde a ningún valor del cuadro 1. Se diagnosticó entonces un cambio de sitio.

Una visita de campo permitió corroborar la fecha del cambio, pero el observador nos indicó un cambio de probeta.

Esas dos estaciones tampoco tienen coeficientes de correlación muy significativos (0,70 y 0,76 para los valores corregidos), lo que puede explicar el error de diagnóstico.

En los dos casos anteriores, fue necesario comprobar las estaciones con relación a todas las otras circundantes para acertar el diagnóstico, lo que constituye un serio limitante de esa metodología.

5.4. Conclusiones sobre los gráficos de dobles masas

Para utilizar este método se debe, en primer lugar, determinar empíricamente grupos de estaciones que pertenezcan a la misma zona climática, teniendo en cuenta la distancia entre ellas, los efectos de relieve, el tipo de régimen intra-anual, etc.

Posteriormente, se compara cada una con todas las demás del grupo, lo que significa un número de dobles masas igual a $m(m-1)/2$ para un grupo de m estaciones. En esas condiciones, el tratamiento de la pluviometría de una región o de un país exige la manipulación de una cantidad considerable de información, posible únicamente mediante automatización de las operaciones de tratamiento.

El programa MASA, utilizado en el proyecto, constituye un buen ejemplo de trazado automático de gráficos. Su utilización facilita considerablemente las comparaciones entre estaciones suprimiendo la mayor parte de las operaciones manuales y aportando una valiosa ayuda en el trabajo de interpretación.

La sola figura 2 no permite tampoco detectar de donde vienen los errores que aparecen en los gráficos con datos originales. Los cambios de pendiente pueden provenir tanto de Mira como de Cahuasquí o de Los Cerros (Montecristi) como de Camarón (Manabí). Varias pruebas serán necesarias para descubrir de donde vienen los problemas.

La mejor manera de utilizar el método de dobles masas es seleccionar, en la zona climática considerada, algunas estaciones con datos de buena calidad y de largo período (estaciones de base). Comparándolas con las estaciones vecinas y corrigiendo cada vez que aparezcan cambios de pendientes, es factible constituir otras estaciones de base y avanzar así lentamente, comprobando cada vez los diagnósticos emitidos con todas las estaciones de base.

Operando minuciosamente y de esa manera, es posible llegar a diagnósticos precisos y comprobados. La metodología es usada en varios continentes en donde se ha probado su eficiencia.

Sin embargo, la realidad es más compleja. No siempre es posible encontrar de antemano estaciones de base y, en tal caso, la interferencia entre errores provenientes de estaciones diferentes dificulta en gran medida la elección de criterios de decisión realmente objetivos.

Además, es frecuente disponer de estaciones de base que tengan períodos de observación incompletos, justamente en los años interesantes de comparar.

Se percibe entonces, el interés de elaborar un vector de referencia:

- homogéneo,
- que cubra todo el período considerado,
- automáticamente,
- en base a los datos existentes.

Con un vector elaborado en condiciones óptimas, su simple comparación con una estación permitiría concluir sin tener que pasar por toda una serie de deducciones cuya práctica conduce rápidamente a confusión cuando el número de estaciones por tratar aumenta.

6. ELABORACION DE UN VECTOR DE REFERENCIA

6.1 Principal interés

Si Z designa al vector de referencia representativo de la tendencia climática regional descrita en el punto 5.2 y en la cual se encuentran las estaciones a verificarse, dicho vector no tendrá fluctuaciones aleatorias y su relación con la tendencia podrá escribirse mediante la siguiente ecuación:

$$(8) \quad z_i = z_M + \sigma_z \cdot w_i$$

en donde z_i representa los valores anuales, z_M el promedio, y σ_z la desviación estándar.

Acumulando las variables z_i hasta el año n y dividiendo para el promedio, la ecuación anterior se escribe:

$$(9) \quad Z_n = n + C_z \cdot \sum w_i$$

en donde C_z corresponde al coeficiente de variación de z_i .

Comparando este vector con una estación X que pertenezca a la misma zona climática y eliminando n entre las ecuaciones (5) y (9), obtenemos:

$$(10) \quad X_n = Z_n + \sum w_i \cdot (r_x \cdot C_x - C_z) + C_x \cdot \sum u_i \cdot \sqrt{1 - r_x^2}$$

o para simplificar:

$$(11) \quad X_n = Z_n + A + B$$

Como lo manifestamos en el punto 5.2, en una misma zona climática, el término A tiende hacia cero, mientras el término B , del cual depende la magnitud de la oscilación de las dobles masas, tiene como varianza:

$$(12) \quad \sigma_v^2 = C_x^2 + C_z^2 - 2 \cdot C_x \cdot C_z \cdot r_x$$

relación en la cual C_x^2 y C_z^2 corresponden al cuadrado de los coeficientes de variación C_x y C_z .

Así, tenemos:

$$(16) \quad A = B + E$$

B representa la información teórica deseada.

6.4 Análisis de algunas posibilidades

Una posibilidad es calcular, para cada año, el valor del vector utilizando la media aritmética de todos los totales observados durante este año.

En caso de homogeneidad previa de todos los datos y de distribución normal, la media aritmética corresponde efectivamente a un parámetro significativo de tendencia central.

Lastimosamente, las variables se alejan frecuentemente de una distribución normal:

- por el escaso número de datos que genera distribuciones asimétricas;
- por la existencia de series heterogéneas regidas por distribuciones plurimodales;
- por la presencia de valores anormales que vienen a alterar el cálculo de la media (« contaminación de la media por los valores extremos »).

Una estimación valiosa puede obtenerse a partir de las medianas, pero los cálculos se vuelven problemáticos cuando las series tienen muchos años sin observación.

En base a las consideraciones anteriores, G. Hiez [6] emprendió la elaboración de un vector regional tomando en cuenta los valores modales. Dicho vector se utiliza en Brasil en donde está dando buenos resultados. Desafortunadamente, el algoritmo es bastante largo en tiempo de cálculo y su aplicación está reservada, por el momento, a grandes computadoras.

Explorando en una dirección diferente, otro investigador (Y. Brunet-Moret [2],[3], [4]) diseñó un « vector de los índices anuales » (o vector YBM) que estima, para cada año, la media más probable.

El proyecto Plan Nacional de Riego utiliza este último vector. A continuación presentamos su teoría y sus resultados.

Cabe señalar que los dos vectores han sido probados en varios países, dando resultados similares, pero nunca han sido aplicados en un país de relieve tan contrastado como lo es el Ecuador.

(7) PRESENTACION DEL VECTOR YBM

7.1 Hipótesis de base

Si **Z** es el vector de la zona climática estudiada, según las ecuaciones (1) y (14), para el año *i*, la precipitación anual $x_{\theta i}$ de la estación θ (de las μ existentes) está vinculada al índice z_i del vector por la relación:

$$(17) \quad x_{\theta i} / x_{M\theta} = z_i + e_{\theta i}$$

$x_{M\theta}$ = esperanza matemática de X_{θ} en su universo

$e_{\theta i}$ = variable aleatoria independiente de z_i

Las variables E_{θ} (de α a μ) son independientes entre ellas.

Como se manifestó inicialmente, para una estación θ , la serie cronológica de $x_{\theta i}$ es de carácter aleatorio sencillo, alterado por efectos de tendencia, persistencia y pseudo-cíclicos.

Dichos efectos están representados ya en el vector; la serie cronológica de los $e_{\theta i}$ es entonces de carácter aleatorio sencillo.

Cabe recordar que en la zona climática considerada existe la hipótesis de que los totales pluviométricos anuales son pseudo-proporcionales.

Todo lo anterior (hipótesis y relación) implica que:

- la esperanza matemática de $E_{\theta} = 0$ en su universo;
- la esperanza matemática de $Z = 1$ en su universo;
- las varianzas de $E_{\alpha}, \dots, E_{\mu}$ son iguales entre ellas y a $\text{Var } E$

Para cada año i , los $e_{\alpha i}, \dots, e_{\mu i}$ no son independientes entre ellos (proximidad geográfica), pero se admite que en sus universos la suma de las covarianzas es igual a cero:

$$\sum \text{cov ar}(E_{\theta}, E_{\delta}) = 0, \quad \theta \neq \delta$$

Esto implica que los coeficientes de correlación entre estaciones no son iguales pero tienen un valor promedio de:

$$(18) \quad \text{CORXYM} = \text{Var } Z / (\text{Var } Z + \text{Var } E)$$

Con esas hipótesis, el coeficiente de correlación entre el vector y una estación cualquiera tendrá en los universos un valor igual a:

$$(19) \quad \text{CORXVM} = \sqrt{\text{CORXYM}}$$

Siendo este valor superior a CORXYM , es comprensible que las dobles masas entre vector y estación sean en general más precisas que las dobles masas entre estaciones.

Lo anterior implica también que el valor del coeficiente de variación de cada estación en su universo sea igual a:

$$(20) \quad C_x = \sqrt{\text{Var } Z + \text{Var } E}$$

7.2 Algoritmo de cálculo

El programa examina los totales anuales de las estaciones seleccionadas a partir de un año de principio (ID) y hasta un año final (IF) y escoge únicamente las estaciones que poseen por lo menos **MINIAÑO** años en el período considerado, así como los años del período en donde existen por lo menos **MINIEST** totales pluviométricos.

Las variables ID, IF, **MINIAÑO** y **MINIEST** son variables de entrada, definidas por el utilizador.

Los valores seleccionados conforman la matriz de observación **A** (n, μ) de las precipitaciones anuales de n estaciones en n años.

Dicha matriz es generalmente incompleta (observaciones faltantes) y constituye la base de cálculo del vector.

En un espacio de μ dimensiones, la ecuación:

$$(21) \quad X_{\alpha} / x_{M\alpha} + X_{\beta} / x_{M\beta} + \dots + X_{\mu} / x_{M\mu} - \mu_{Z1} = 0$$

representa un hiperplano perpendicular a la recta:

$$(22) \quad X_{\alpha} / x_{M\alpha} = X_{\beta} / x_{M\beta} = \dots = X_{\mu} / x_{M\mu}$$

El cálculo consiste en determinar los valores $x_{M\theta}$ para cada estación y los valores de z_i para cada año, para que los hiperplanos representativos de las observaciones de cada año sean paralelos entre ellos y perpendiculares a la recta (22).

Según las hipótesis anteriores aplicadas a la muestra (promedio de $z = 1$, promedio de $E = 0$, $\text{Var } \theta = \text{Var } E$), se pueden utilizar los mínimos cuadrados para minimizar la expresión:

$$(23) \quad \sum \sum (x_{\theta i} / x_{M\theta} - z_i)^2$$

para determinar $x_{M\theta}, \dots, x_{M\mu}, z_1, \dots, z_n$.

Derivando dicha expresión según z_i , si existen μ_1 estaciones observadas en el año i , obtenemos n ecuaciones:

$$(24) \quad z_i = 1 / \mu_1 \sum (x_{\theta i} / x_{M\theta})$$

Derivando la expresión (23) según $1/x_{M\theta}$, si existen n_1 años observados en esta estación θ , obtenemos μ ecuaciones:

$$(25) \quad \sum x_{\theta i} (x_{\theta i} / x_{M\theta} - z_i) = 0$$

En realidad, teniendo en cuenta que existen datos faltantes, se debe reemplazar z_i en la expresión (23) por su valor encontrado en la ecuación (24) antes de derivar según $1/x_{M\theta}$ para obtener un sistema de μ ecuaciones con μ incógnitas ($1/x_{M\theta}$):

$$(26) \quad f(\alpha, \theta)(1/x_{M\alpha}) + \dots + f(\mu, \theta)(1/x_{M\mu}) = 0$$

Este sistema lineal es homogéneo y tiene una solución banal:

$$1/x_{M\alpha} = \dots = 1/x_{M\mu} = 0$$

Para su resolución, debe ser reducido a $(\mu - 1)$ ecuaciones lineales de $(\mu - 1)$ incógnitas y con el segundo miembro diferente de cero.

Fijando $1/x_{M\mu} = 1$, se calculan entonces, por los mínimos cuadrados (regresión múltiple), los $(\mu - 1)$ valores de $1/x_{M\theta}$ que mejor satisfacen las μ relaciones entre las cantidades f de la ecuación (26).

Se conocen así los valores $C/x_{M\alpha}, \dots, C/x_{M\mu}$, siendo C un coeficiente multiplicador desconocido, pero igual para todas las variables.

Con las ecuaciones (24) se calculan los valores de z_i . Formulando la hipótesis de que el promedio de los z_i es igual a 1, se calcula el coeficiente C , pudiéndose así determinar exactamente los valores $x_{M\alpha}, \dots, x_{M\mu}$.

El cálculo anterior muestra que:

- para cada año i , $z_i - 1 / \mu_1 \sum x_{\theta i} / x_{M\theta} = 0$ (27)
- para cada estación θ , $\sum (x_{\theta i} / x_{M\theta} - z_i) \neq 0$ (28)
- $\sum \sum (x_{\theta i} / x_{M\theta} - z_i) = 0$ (29)

APLICACION A CASOS CONCRETOS

El algoritmo anterior da el mismo peso a cada uno de los valores observados al igual que a los valores alterados que se quieren detectar y que pueden influir en los valores de x_{M0} y z_1 .

Conviene entonces formular una última hipótesis admitiendo que en cada año y en cada estación los valores erróneos son la minoría.

El cálculo se realiza según el siguiente proceso:

- * Una primera iteración se efectúa con los valores reales de la matriz **A** y se calcula una primera serie de valores para los x_{M0} y z_1 .

Para cada año y para cada estación se compara x_{0i}/x_{M0} con z_1 . Si $x_{0i}/x_{M0}/z_1$ no está comprendido entre un intervalo dado, x_{0i} será reemplazado por $x_{M0} \cdot z_1$.

- * Los nuevos valores de la matriz **A** se utilizan para una segunda iteración, la cual calcula nuevos valores de x_{M0} y z_1 que se comparan nuevamente con el intervalo y se reemplazan de ser necesario según el procedimiento anterior.
- * Se procede a una serie de iteraciones hasta que todos los valores de $x_{0i}/x_{M0}/z_1$ se encuentren al interior del intervalo, lo que ocurre rápidamente.

Es lógico que los valores que delimitan el intervalo dependan de la varianza de **E** de la cual tenemos una estimación (por exceso si la homogeneidad espacial no es realizada). En la realidad, es difícil tener en cuenta esa observación cuando existen series heterogéneas.

Después de varias tentativas, tanto en el callejón interandino como en la Costa, se resolvió tomar 0,7 y 1,5 como valores que delimitan el intervalo de comparación.

Esas cantidades son similares a $2/\pi$ y a $\pi/2$, valores debidos a los errores más frecuentes entre probetas y pluviómetros (cuadro nº 1).

7.3 Test de sensibilidad

Consideramos un grupo de 12 estaciones ficticias observadas durante 20 años (1966-1985 por ejemplo) y cuyos totales anuales varían entre 999 y 1.001 mm (casi siempre iguales a 1.000 mm).

Los datos de 3 estaciones en un período de 10 años (1966-1975) son multiplicados por 5.

Un vector elaborado a partir de la media aritmética sería alterado y tendría 2 períodos cuyas pendientes diferentes tendrían una relación igual a 2.

Con el vector, la relación es de 1,25, lo que demuestra su poca sensibilidad a confusiones de aparatos o a cambios de sitio.

En esas mismas estaciones, se multiplica por 13 la precipitación de 1985 en una estación.

Para ese año, el valor de la media alcanza los 2.000 mm, mientras que el valor del vector es de 1.025 mm, lo que pone en evidencia su casi insensibilidad a valores extremos como errores de teclado, etc.

En los dos casos anteriores (un tanto extremos), la comparación de las estaciones con la media no hubiera permitido detectar los errores, al estar ella misma alterada.

8. RESULTADOS OBTENIDOS

Desde hace varios meses, el equipo de trabajo está efectuando el tratamiento de la información de más de 700 estaciones de la Sierra y de la Costa, aplicando la metodología anterior.

La cuenca del río Mira está terminada y los diagnósticos emitidos han sido comprobados con verificaciones de campo y revisión de los historiales en el INAMHI.

En la parte sur del país y en las cuencas del Guayas y del Pastaza, el tratamiento ha terminado y sólo resta verificar los diagnósticos.

La homogeneización está bien avanzada en Manabí y en la cuenca del río Esmeraldas.

Solamente queda por emprender los análisis en las estaciones ubicadas en las cuencas de Cañar y Zapotal.

El concepto de pseudo-proporcionalidad ha sido verificado en todas partes, aunque las zonas climáticas encontradas en la Sierra son más reducidas y sus límites más difíciles de trazar.

Por lo tanto, no existen tantos micro-climas como se podía esperar y las estaciones obedecen a efectos climáticos similares cuyo impacto en las precipitaciones varía según la posición de la estación en relación con el relieve (efectos de abrigo, etc.).

Más vale entonces hablar de topo-climas.

El vector detecta también los errores con buena confiabilidad como lo observamos en la figura 4 que representa las estaciones El Angel y Zuleta ya presentadas en la prueba de dobles masas (figura 3).

Los coeficientes de correlación han aumentado significativamente, lo que permite acertar el diagnóstico por visualización de un solo gráfico.

El cambio de sitio de la estación El Angel aparece ahora entre 1969 y 1970 (conforme al historial), y la relación de pendientes en el caso de Zuleta es de 1,57 confirmando así el cambio de probeta realizado por el observador en 1972.

Las verificaciones han permitido confirmar todos los diagnósticos de la cuenca del Mira para la cual se obtuvieron los siguientes resultados:

- 58 estaciones analizadas
- 14 con cambios de sitio o inadecuación de aparatos
- 17 con años anormales

es decir casi un 50 % de estaciones homogéneas, lo que corresponde a los resultados obtenidos en países vecinos.

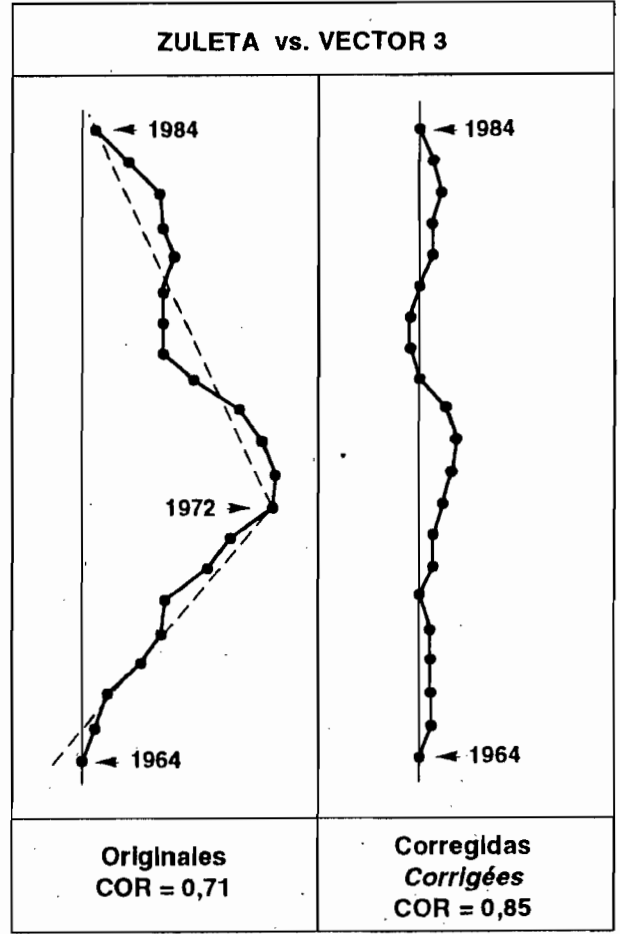
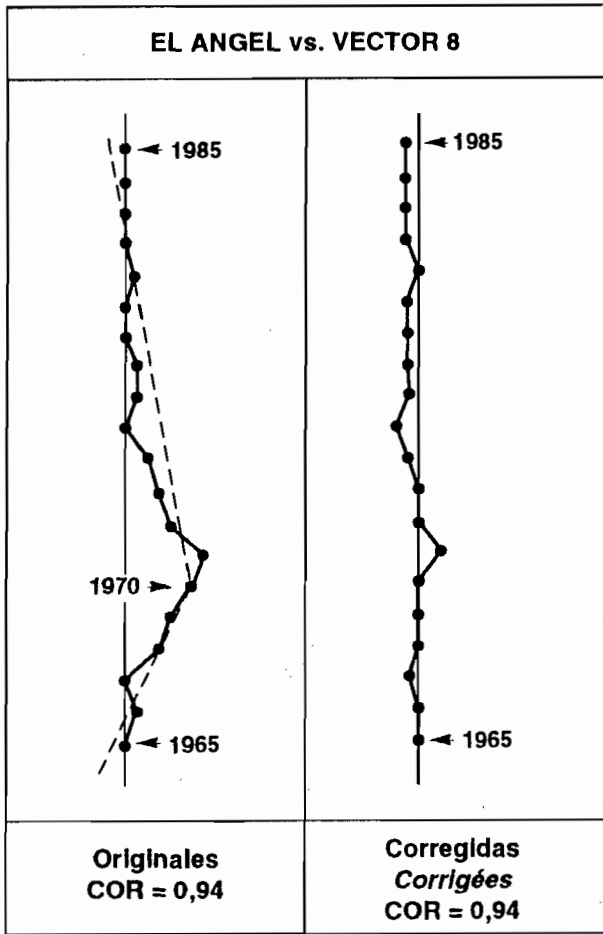


Figura 4

$E = F$
 $R = F \cdot P$

B - REGIONALIZACION

1. DEFINICION DEL PROBLEMA

La regionalización pluviométrica, su concepción y la precisión de su cálculo dependen en gran medida de los objetivos perseguidos.

El Plan Nacional de Riego necesita establecer, en cualquier parte del país, balances entre necesidades de agua y recursos hídricos, los cuales tendrán que calcularse mes a mes y año por año para tener en cuenta los riesgos (desde el punto de vista estadístico) que deberán asumir los perímetros regables en cuanto a disponibilidad de agua.

- En el caso del cálculo de las necesidades de agua, se requiere entonces disponer de una serie cronológica mensual y anual homogénea, observada en un período suficientemente largo y cuya precipitación anual promedio sea igual a la del perímetro considerado.

La utilización de un vector de referencia no permitirá obtener datos exactos (ausencia de fluctuaciones aleatorias en el sitio considerado), pero permitirá estimar los parámetros estadísticos correctos.

- Por otra parte, el cálculo de la lluvia promedio de una cuenca vertiente integra las precipitaciones puntuales de cada una de sus partes y por lo tanto elimina o reduce la influencia de las fluctuaciones aleatorias debidas a cada medición. En ese caso, la estimación del recurso hídrico mediante la relación lluvia-caudal se aproximará más a la realidad si utiliza también un vector de referencia que no contenga esos efectos.

Lo anterior demuestra que los balances hídricos en un punto dado pueden establecerse de mejor manera a partir de un vector de referencia, representativo de la región del sitio considerado (o de su cuenca vertiente), desde el punto de vista tanto cronológico (regionalización cronológica) como de su repartición intra-anual (regionalización de régimen), siempre y cuando se conozca la precipitación anual promedio del punto de cálculo o de su cuenca.

Veamos como el vector YBM puede resolver esos problemas.

2. TRAZADO DE ISOYETAS

Para ser efectivas, las isoyetas deben trazarse en base a estaciones cuyos totales anuales hayan sido calculados en un período común suficientemente largo para ser representativo.

En teoría, el vector YBM se elabora según la esperanza matemática de las precipitaciones anuales de cada estación.

En la práctica, el programa de cálculo estima cada esperanza matemática en base a la organización interna de la tendencia climática regional.

No se necesita entonces rellenar los años faltantes de cada estación en un período común ni calcular la representatividad de tal período. Además, las relaciones entre lluvia y altitud son normalmente más precisas al interior de cada zona climática encontrada.

3. REGIONALIZACION CRONOLOGICA

La metodología del vector al igual que la de las dobles masas depende de la existencia de zonas climáticas homogéneas.

Los límites de las zonas 4 y 8 no están completamente definidos sencillamente por falta de estaciones cercanas.

Se han obtenido resultados similares en las otras regiones, y es posible también detectar las partes en las que la red pluviométrica es insuficiente.

4. REGIONALIZACION DE REGIMEN

Para agrupar estaciones según su régimen, es posible razonar, como usualmente, en promedios mensuales.

Lastimosamente, el régimen de una estación varía según el año, y existen en la Sierra estaciones de régimen promedio interandino mientras tienen algunos años de régimen pseudo-amazónico, como el caso del año 1975 en la estación Ibarra-Aeropuerto (figura 6). Así, aunque la comparación en régimen promedio es válida, no es suficiente, y es necesario, para no perder información, considerar los datos mensuales como una sola serie cronológica.

Desde ese punto de vista, las hipótesis y ecuaciones que rigen la elaboración del vector no impiden considerar una serie mensual.

El concepto de pseudo-proporcionalidad será más difícil de encontrar, las fluctuaciones aleatorias serán mayores y los ajustes a valores anuales tal vez delicados. Sin embargo, el grupo de trabajo está realizando ensayos en ese sentido, y esperamos que resultados futuros vengán a confirmar este razonamiento.

