

Resultados de herramienta informática para recuperación de registro faltante en series temporales

Enrique A. Sánchez Camacho

Resumen

Hace algunos meses, en este espacio se me dio la oportunidad de exponer el contenido y características de una herramienta informática construida en *Fortran95*, cuyo fin, en términos generales, es recuperar el registro faltante de series temporales (Sánchez, 2013). Los motivos, la teoría que sustenta a la herramienta y la descripción de la misma se hicieron en el mencionado escrito, por lo que se omitirá tal información en el presente trabajo.

Alcanzada la meta, esta se vuelve punto de partida. Con intención de probar de manera práctica lo expuesto hace unos meses, ahora se presentan algunos resultados obtenidos con la utilización de la herramienta informática referida. Las metodologías contenidas en ella son el método de proporción normal (Paulhus y Kohler, 1952), y el método de inverso de la distancia (ASCE, 1996).

Los registros recuperados por ambas metodologías se comparan con el registro original disponible –antes de recuperar su parte faltante–, a través de sus promedios (índice de tendencia central) y desviación estándar (índice de dispersión); lo obtenido para el caso empleado es muy satisfactorio, como se muestra en los siguientes apartados.

Introducción

“La teoría pretende ser la práctica sin sudor”. La frase, además de elegante, para quienes nos aventuramos a usar la teoría en la práctica sabemos que su afirmación irrestricta ocurre pocas veces. El punto de partida entonces –en este caso– comienza con la delimitación de una zona de interés: Subregión Huixtla y otros, conformada por tres subcuencas: Laguna del Viejo y Tembladeras, con 892.56 km²; Río Despoblado, con 605.11 km² y Río Huixtla con 828.97 km², en el estado de Chiapas. La figura 1 muestra la ubicación geográfica de la zona de estudio.

Dentro de las tres subcuencas, se encuentran cinco estaciones hidrométricas que aforan las cuatro corrientes principales de la subregión Huixtla y otros. Para lo aquí presentado, usaremos cuatro de las cinco estaciones hidrométricas existentes. La estación 23031 Compoapa se encuentra en uno de los afluentes del río Despoblado, pero su registro tiene faltantes importantes, circunstancia que llevó a omitir el registro de tal estación para este ejercicio. La posición relativa de las estaciones hidrométricas y los principales ríos se muestran en la figura 2.

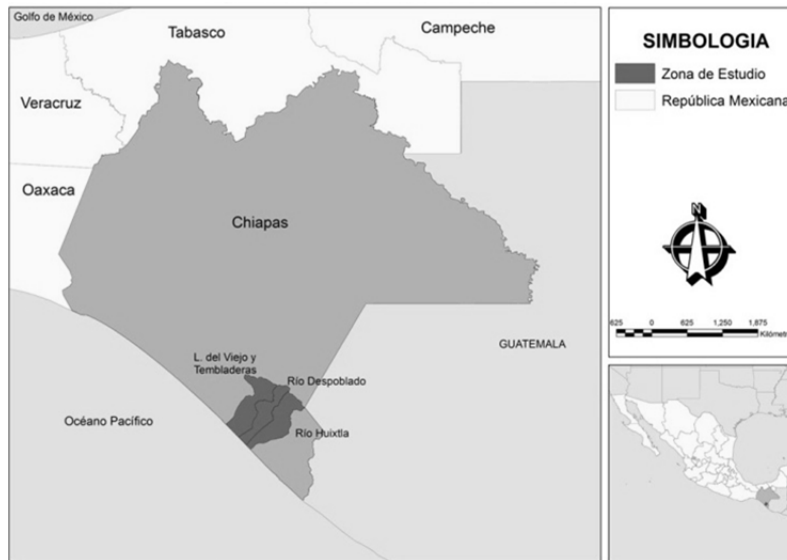


Figura 1. Ubicación geográfica de área de estudio.

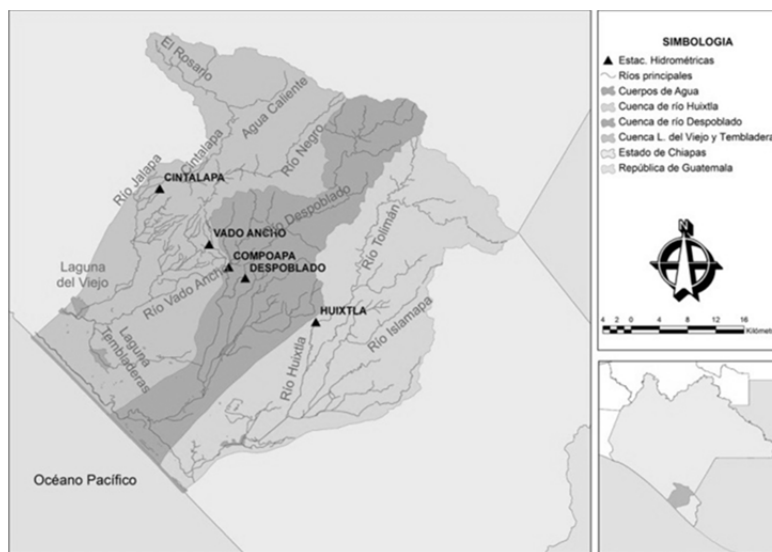


Figura 2. Posición relativa de subcuencas y estaciones hidrométricas.

En la tabla 1 se muestran las condiciones del registro para el periodo 1970 a 2001 –elegido para el ejercicio– en cada estación hidrométrica, luego de haber revisado su disponibilidad de información de manera particular.

Tabla 1. Estado del registro de las estaciones hidrométricas.

| Clave | Nombre corriente | Inicio Año/Mes | Fin Año/Mes | % Información |
|-------|------------------|-------------------|----------------|---------------|
| 23016 | Cintalapa | 1970/ene | 2001/oct | 88.8 |
| 23032 | Vado Ancho | 1974/ene | 1983/nov | 30.7 |

| | | | | |
|---|------------|----------|----------|------|
| 23019 | Despoblado | 1970/ene | 2001/oct | 95.8 |
| 23008 | Huixtla | 1970/ene | 2001/oct | 84.1 |
| % Información: cantidad de información que tiene el registro de estación para el periodo considerado. | | | | |

En la misma tabla puede observarse que la estación hidrométrica 23032 (Vado Ancho) es la que tiene el registro más incompleto para el periodo seleccionado, lo cual lo convierte en el objeto del ejercicio que a continuación se presenta. La información utilizada y mostrada de manera sintética aquí proviene del Banco Nacional de Aguas Superficiales (Conagua, 2012).

La recuperación del registro de la hidrométrica 23032 (Vado Ancho) se realizó con el apoyo de los registros de las tres restantes estaciones hidrométricas, empleando los métodos de proporción normal e inverso de la distancia, de modo que se obtuvieron dos registros para el periodo temporal señalado en tal estación hidrométrica.

Resultados de la herramienta informática

Tras recuperar el registro de la estación hidrométrica Vado Ancho 23032 por las dos metodologías señaladas, se realizó un análisis estadístico básico. El primero consistió en calcular el volumen promedio para cada mes en el periodo 1970-2001, sobre los resultados de cada una de las metodologías empleadas.

Al mismo tiempo, este cálculo de volumen promedio se obtuvo utilizando el registro de la estación hidrométrica Vado Ancho 23032 con sus datos originales solamente, con el fin de tener un referente con el cual comparar lo calculado.

El resultado de lo antes descrito se muestra en la figura 3. Lo más evidente es la forma de las tres curvas, bastante próximas entre sí. Luego, si consideramos como referente la línea del registro original, la curva que se le aproxima más es la obtenida por el método de proporción normal.

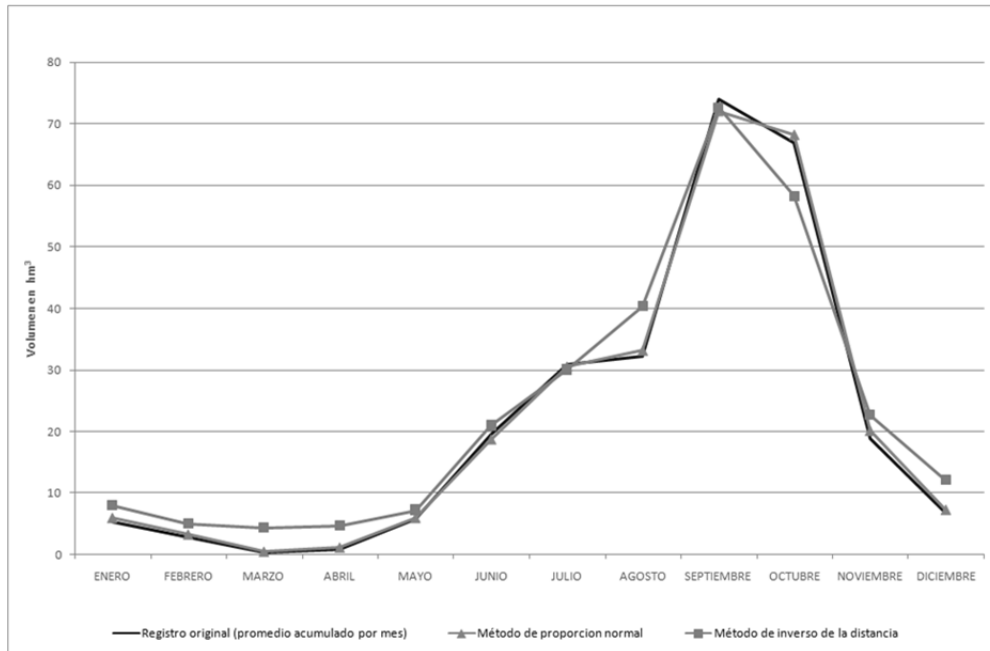


Figura 3. Comparación del promedio en volumen acumulado promedio de registro recuperado y registro original.

A continuación, se calculó la desviación estándar (índice de dispersión). Como en el caso anterior, se realizó sobre el promedio mensual del periodo 1970-2001 para los dos registros calculados y el registro original, antes de recuperar su parte faltante. El comparativo se muestra en la figura 4.

En este caso, la dispersión del registro original es mayor en los meses de mayo a octubre que la obtenida en los dos registros recuperados, aunque de noviembre a febrero sus valores están por debajo de los que tienen los registros recuperados.

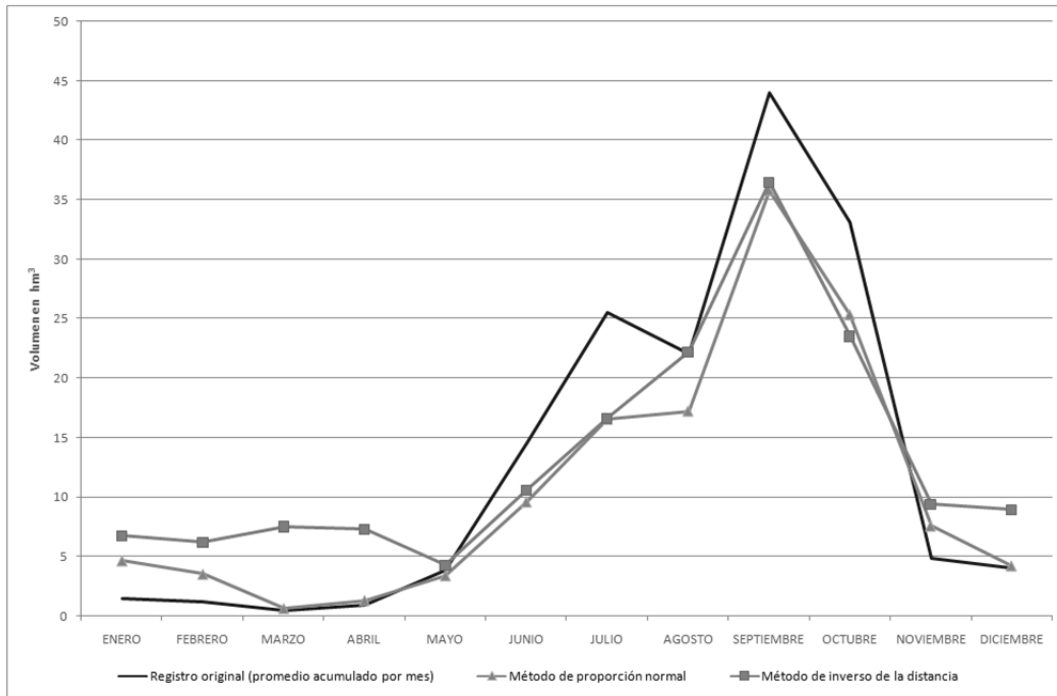


Figura 4. Comparación de desviación estándar en volumen acumulado mensual promedio de registro recuperado y registro original.

CONCLUSIONES

Con lo hasta aquí mostrado, los resultados que ofrece la herramienta informática para recuperar registro faltante son bastante satisfactorios, como se afirmó al comienzo de este escrito. Los resultados, hablando coloquialmente, “están en rango”. Se observa en las figuras de las curvas de promedio y de desviación estándar que se encuentran lo suficientemente cercanas a la forma de la curva del registro original.

Desde el punto de vista numérico, las diferencias mayores entre los estadísticos del registro original y los calculados se encuentran en la curva de la desviación estándar para el método de inverso de la distancia, aunque existe una variación de signos en las diferencias, mes a mes.

Finalmente, un comparativo más próximo que los anteriores ocurre en la figura 5. En ella se ha dibujado el registro mes a mes. El registro original de la estación hidrométrica Vado Ancho 23032 se identifica con la característica en la que sólo aparece dibujado un grupo de datos; en el resto del gráfico, se observan los dos grupos de datos calculados, donde de forma notable coinciden en valores. Este último hecho nos indica una respuesta satisfactoria con ambas metodologías.

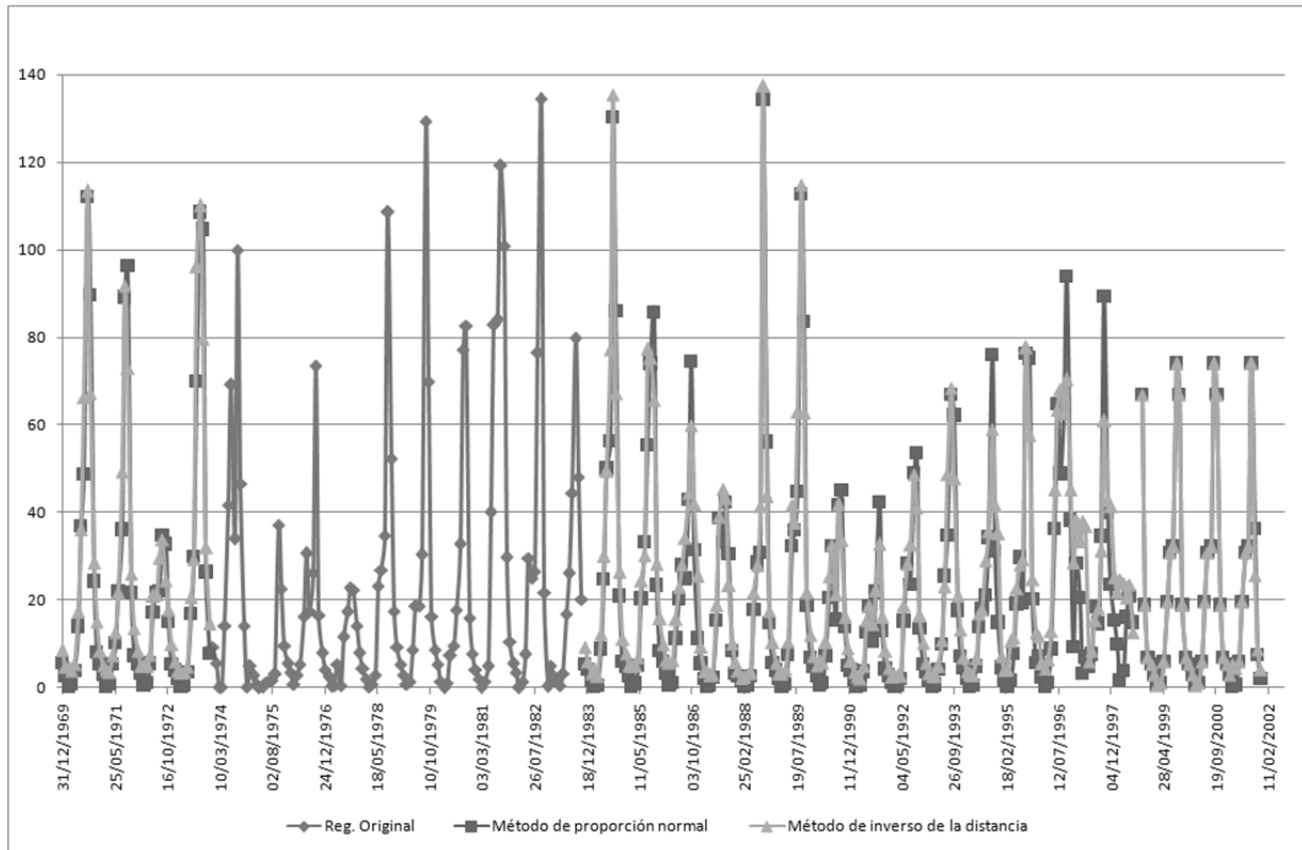


Figura 5. Comparación de los registros calculados, con la inclusión del registro original de la estación hidrométrica 23032.

Referencias bibliográficas

ASCE, American Society of Civil Engineers (1996), *Hydrology Handbook*, ASCE, New York, pp. 5-74.

Conagua, Comisión Nacional del Agua (2012), *Banco Nacional de Datos de Aguas Superficiales (BANDAS)*, Consulta de datos hidrométricos, de sedimentos y vasos, <http://www.conagua.gob.mx/CONAGUA07/Contenido/Documentos/Portada%20BANDAS.htm> Consulta: junio 4, 2012.

Paulhus, J. L. H. y M. A. Kohler (1952), "Interpolation of Missing Precipitation Records", *Monthly Weather Review*, vol. 80, N° 8, pp. 129-33.

Sánchez Camacho, E. A. (2013), "Relleno de registro faltante en estaciones climatológicas", *Boletín IMTAnet*, N° 17, 2013.