

# Validación del modelo *ClimGen* en la estimación de variables de clima ante escenarios de datos faltantes con fines de modelación de procesos

• Gerardo Esquivel\* •

*Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, México*

\*Autor de correspondencia

• Julián Cerano •

*Universidad Nacional Autónoma de México*

• Ignacio Sánchez •

*Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, México*

• Armando López •

*Universidad Autónoma Chapingo, México*

• Oscar G. Gutiérrez •

*Universidad Estatal de Sonora, México*

## Resumen

Esquivel, G., Cerano, J., Sánchez, I., López, A., & Gutiérrez, O. G. (julio-agosto, 2015). Validación del modelo *ClimGen* en la estimación de variables de clima ante escenarios de datos faltantes con fines de modelación de procesos. *Tecnología y Ciencias del Agua*, 6(4), 117-130.

Diversos modelos hidrológicos y ambientales con frecuencia requieren información meteorológica, como datos de entrada a intervalos de tiempo variable, que a menudo no están disponibles en los sitios de interés. En la mayoría de las estaciones meteorológicas, el periodo de registro de datos es a menudo insuficiente para permitir una buena modelación de procesos, por lo tanto existe una seria restricción en su aplicación si no se dispone de manera directa de los datos meteorológicos. En este estudio, el generador climático *ClimGen* fue evaluado para la estimación de datos faltantes de temperatura y precipitación para tres sitios con baja, media y alta ocurrencia de precipitación. Se realizó la parametrización y cálculo de datos faltantes para la estación meteorológica Riito, que representa la condición seca, obteniendo valores bajo el criterio de  $r^2$  para la variable temperatura máxima de  $r^2 = 0.96$ , temperatura mínima  $r^2 = 0.95$  y precipitación  $r^2 = 0.98$ . Las estaciones Tepehuanes y El Tarahumar, que representan la condición media de humedad, obtuvieron valores para la variable temperatura máxima de  $r^2 = 0.98$ ; temperatura mínima  $r^2 = 0.90$  y  $r^2 = 0.99$ ; y precipitación  $r^2 = 0.96$  y  $r^2 = 0.93$ , respectivamente. Por último, la estación Francisco Rueda, que representa la condición húmeda, obtuvo  $r^2 = 0.96$  para temperatura máxima;  $r^2 = 0.98$  para temperatura mínima, y  $r^2 = 0.97$  para precipitación. Los resultados indican que los datos estimados por el generador climático son representativos de los datos históricos del clima en los sitios de estudio.

**Palabras clave:** modelación de procesos, generadores climáticos, precipitación, temperatura, *ClimGen*.

## Abstract

Esquivel, G., Cerano, J., Sánchez, I., López, A., & Gutiérrez, O. G. (July-August, 2015). Validation of the *ClimGen* Model to Estimate Climate Variables when Lacking Data for Modeling Processes. *Water Technology and Sciences (in Spanish)*, 6(4), 117-130.

Many hydrological and environmental models often require meteorological information corresponding to different time intervals as input data. This information is often not available at the sites of interest. At most weather stations, data registry periods are frequently insufficient for a good modeling of processes. A series of restrictions exist in their application when meteorological data is not directly available. The present study evaluated the use of the *ClimGen* weather generator to estimate missing temperature and rainfall data for three sites with low, medium and high rainfall. The parameterization and calculation of the missing data performed for the Riito weather station, which represents dry conditions, resulted in  $r^2$  values for maximum temperature of  $r^2 = 0.96$ , minimum temperature  $r^2 = 0.95$  and rainfall  $r^2 = 0.98$ . The Tepehuanes and El Tarahumar stations represent medium rainfall conditions, and resulted in  $r^2$  values for maximum temperature of  $r^2 = 0.98$ , minimum temperature  $r^2 = 0.90$  and  $r^2 = 0.99$ , and rainfall  $r^2 = 0.96$  and  $r^2 = 0.93$ , respectively. Lastly, The Francisco Rueda stations represented high rainfall conditions and resulted in  $r^2$  values for maximum temperature of  $r^2 = 0.96$ , minimum temperature  $r^2 = 0.98$  and rainfall  $r^2 = 0.97$ . The results indicate that the data estimated by the weather generator are representative of historical climate data at the study sites.

**Keywords:** Process modeling, weather generators, precipitation, temperature, *ClimGen*.

Recibido: 12/05/2014

Aceptado: 24/04/2015

## Introducción

Las variables climáticas, como precipitación y temperatura, influyen en gran medida en el ciclo hidrológico, y cambios en estas variables podrían afectar los patrones de evaporación y escurrimiento, la cantidad de agua almacenada en glaciares, humedales, lagos y aguas subterráneas (Sarangi & Kumar, 2006). Estas variables son factores determinantes en la producción agrícola, sobre todo en los sistemas de cultivo que se encuentran por lo común en las regiones áridas y tropicales (Tingem, Rivington, Azam-Ali, & Colls, 2008); por ende, esta variabilidad afecta los procesos hidrológicos que pueden generar eventos extremos como inundaciones y/o la ocurrencia de sequías.

En el contexto de la simulación de procesos, los modelos hidrológicos y ambientales se han convertido en herramientas importantes para la planeación, manejo y gestión de recursos naturales. Sin embargo, estos modelos requieren de diferentes datos de entrada (temperatura máxima y mínima, precipitación, radiación solar, velocidad de viento y escurrimiento, entre otros) a intervalos de tiempo variable, que a menudo no están disponibles en los sitios de interés (Safeq & Fares, 2011). En la mayoría de las estaciones meteorológicas, el periodo de registro de datos es a menudo insuficiente para permitir una buena modelación de procesos, por lo tanto existe una seria restricción en la aplicación de modelos si no se dispone de manera directa de los datos meteorológicos (Hoogenboom, 2000). Esta situación resalta la importancia de extender los registros de datos mediante la generación de datos climáticos derivados de las observaciones a corto plazo, utilizando diferentes procedimientos estadísticos.

Los modelos matemáticos conocidos como generadores de tiempo estocástico que simulan series de tiempo de variables climáticas han abordado este problema (Richardson & Wright, 1984). Estos modelos proporcionan datos para incrementar el registro existente de un sitio o a través de la interpolación de los parámetros del modelo proveen información climática en

donde no se dispone de datos medidos. Estos modelos tienen varios componentes interconectados y por lo general simulan múltiples variables utilizando información meteorológica histórica observada como datos de entrada y generan datos meteorológicos estimados, que son estadísticamente similares a los registros climáticos observados (Hoogenboom & Soltani, 2003).

WGEN fue uno de los primeros generadores climáticos desarrollados para su uso con modelos de erosión de suelos y calidad de agua (Richardson & Wright, 1984). Otros generadores de tiempo han sido desarrollados desde entonces. CLIGEN, generador climático incorporado dentro del modelo *Water Erosion Prediction Procedure* (Flanagan & Livingston, 1995) se basa en los métodos de generación climáticos utilizados en WGEN (Nicks, Richardson, & Williams, 1990). Otros generadores climáticos incluyen USCLIMATE (Johnson, Hanson, Hardegree, & Ballard, 1996), CLIMAK (Danuso & Della, 1997) y ClimGen (Stöckle, Campbell, & Nelson, 1999).

Las bases de datos de registros climáticos en México, por lo general carecen de series extensas que permitan analizar la variabilidad climática en una escala de tiempo mayor a los 60 años; además, los registros presentan un alto porcentaje de datos faltantes (IMTA, 2009), representando un mayor problema de confiabilidad para ser empleadas en estudios climáticos, hidrológicos, ambientales y paleoclimáticos. Ante tal problemática, el presente trabajo se plantea como objetivos: 1) evaluar el generador climático *ClimGen* (versión 4.06.06) (Stöckle, Nelson, Donatelli, & Castellví, 2001) para estimar datos faltantes en estaciones meteorológicas, y 2) mejorar y extender los registros meteorológicos de las estaciones.

## Materiales y métodos

### Área de estudio

El presente trabajo se desarrolló para tres sitios de la república mexicana con condiciones medioambientales contrastantes (baja, media y

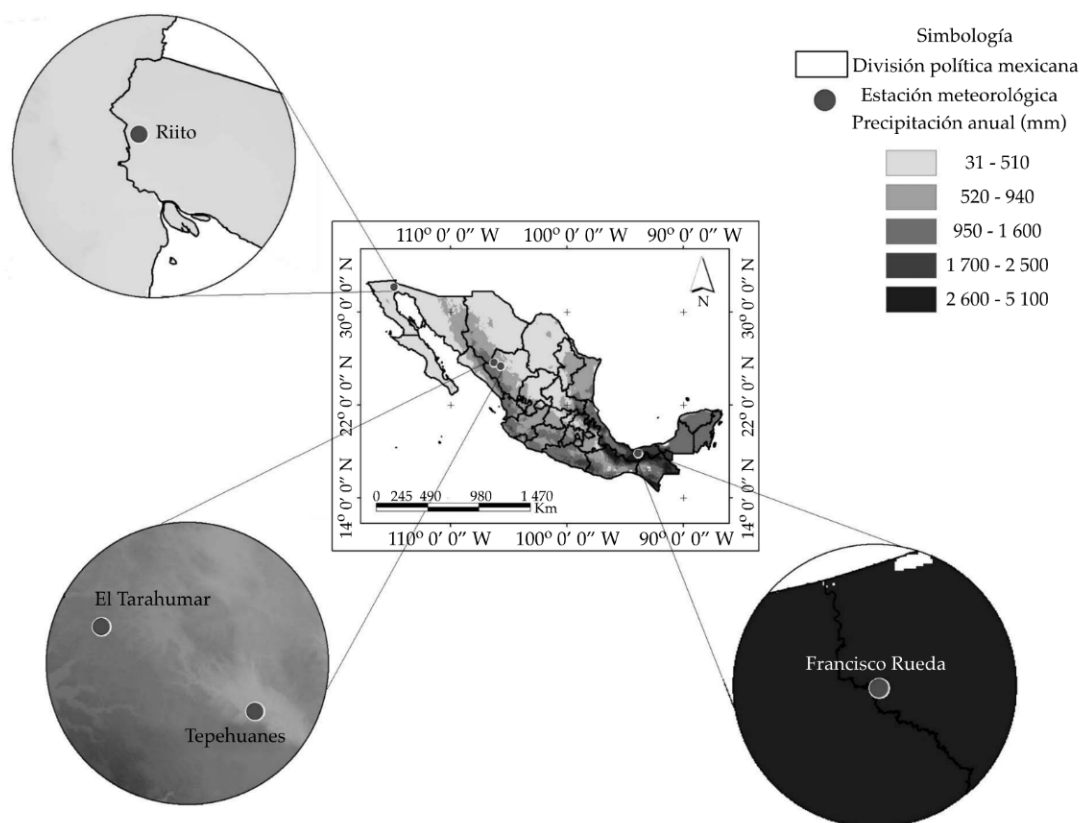


Figura 1. Localización geográfica de las estaciones meteorológicas de estudio.

alta ocurrencia de precipitación) (figura 1). La estación Riito, del municipio de San Luis Río Colorado, en el estado de Sonora, se localiza entre los  $32^{\circ} 08' 00''$  latitud norte y  $114^{\circ} 54' 05''$  longitud oeste, con una precipitación promedio anual de 52.5 mm y  $22.1^{\circ} \text{C}$  de temperatura media anual, a una altitud de 13 metros sobre el nivel del mar (msnm). Las estaciones meteorológicas Tepehuanes y El Tarahumar se enclavan en la Sierra Madre Occidental, en el estado de Durango, entre los  $25^{\circ} 20' 00''$  y  $25^{\circ} 37' 01''$  de latitud norte,  $105^{\circ} 43' 00''$  y  $106^{\circ} 19' 28''$  de longitud oeste, respectivamente; el gradiente altitudinal oscila de los 1 680 a 3 120 msnm, con una precipitación anual de 676 mm y una temperatura media anual de  $13.5^{\circ} \text{C}$ . Por último, la estación Francisco Rueda, en el municipio de Huimanguillo, Tabasco, se ubica en los  $17^{\circ} 50'$

$12''$  de latitud norte y  $93^{\circ} 56' 30''$  de longitud oeste, con una precipitación promedio anual de 2 414 mm,  $26.3^{\circ} \text{C}$  de temperatura media anual y una altitud de siete msnm (IMTA, 2009; Conagua, 2012).

### Información climática

Los registros históricos de clima fueron obtenidos del *Extractor Rápido de Información Climatológica III v. 2.0 (ERIC III)* (IMTA, 2009). Para cada una de las series de datos se realizó un análisis de consistencia de la información, a fin de detectar datos atípicos atribuibles a errores de captura o de otro tipo (p. ej., temperatura máxima con valores de 0; temperatura máxima inferior a la mínima; meses con registros repetidos, etc.), con el objetivo de mejorar la calidad de los datos

de entrada y por consiguiente disminuir errores durante el proceso de validación; asimismo, se obtuvieron los estadísticos descriptivos para cada una de las estaciones de estudio (cuadro 1). El esquema metodológico para descartar inconsistencias para cada variable, temperatura máxima ( $t_{\text{máx}}$ ) temperatura mínima ( $t_{\text{mín}}$ ) y precipitación (PP) consistió en la metodología propuesta por Castro y Carbajal (2010).

### Descripción de ClimGen

*ClimGen* es un generador estocástico de tiempo meteorológico que se utiliza para estimar series de datos a escala diaria de variables como precipitación, temperatura máxima, temperatura mínima, radiación solar, humedad relativa y velocidad de viento, cuyos estadísticos resultan similares a los datos climáticos históricos (Tingem, Rivington, Azam-Ali, & Colls, 2007), pudiéndose parametrizar para cada estación (Safeeq & Fares, 2011). *ClimGen* utiliza la distribución Weibull (Weibull, 1951) para representar la precipitación diaria y la curva de aproxima-

ción constituye una mejora en términos de las series Fourier utilizadas por otros generadores climáticos (p. ej., *CLIGEN*) para simular las variaciones estacionales en los datos de clima.

### Precipitación

Para generar el número de eventos de precipitación se utiliza un modelo de cadenas de Markov de segundo orden (Nicks et al., 1990). La combinación de las probabilidades condicionales para los dos estados de las cadenas de Markov ( $\alpha$ : día húmedo seguido de un día seco;  $\beta$ : día seco seguido de un día húmedo) es calculada de forma individual para cada estación en forma mensual a partir de datos históricos observados (Nicks et al., 1990), como se muestra a continuación:

$$P(W/D) = \alpha$$

$$P(D/D) = 1 - \alpha$$

$$P(D/W) = \beta$$

$$P(W/W) = 1 - \beta$$

Cuadro 1. Estadística descriptiva de las estaciones meteorológicas consideradas en el estudio.

	Estación	Riito	Tepehuanes	El Tarahumar	Francisco Rueda
	Años de registro	1949-2008 (59 años)	1922-1988 (66 años)	1964-2009 (45 años)	1965-2007 (42 años)
Temperatura máxima ( $t_{\text{máx}}$ )	Media	31.2	26.4	19.1	31.6
	Desviación estándar	8.5	5.5	4.6	4.2
	Varianza	72.9	30.4	21.1	17.8
	Máximo	50	40	32.5	45
	Mínimo	8	7.5	4	15
Temperatura mínima ( $t_{\text{mín}}$ )	Media	12.9	7	1	20.7
	Desviación estándar	7.8	6.2	6.7	2.8
	Varianza	61.4	38.9	44.3	8.1
	Máximo	33	20	17	27
	Mínimo	-6.5	-11.2	-19	9
Precipitación (PP)	Media	0.14	1.3	2.5	6.5
	Desviación estándar	1.7	4.3	6.9	15.5
	Varianza	2.9	18.9	47.7	240.1
	Máximo	80	78.5	120	187.5
	Mínimo	0	0	0	0



Donde:

- $P(W/D)$  = es la probabilidad de que ocurra un día con precipitación, dado que el día anterior no hubo lluvia.
- $P(D/D)$  = es la probabilidad de que haya un día sin precipitación, dado que el día anterior tampoco pasó.
- $P(D/W)$  = es la probabilidad de que se tenga un día sin precipitación, dado que el día anterior si llovió.
- $P(W/W)$  = es la probabilidad de que un día llueva, dado que el día anterior también llovió.

*ClimGen* utiliza una distribución Weibull de dos parámetros para calcular la magnitud de la precipitación en días húmedos. La distribución Weibull ha demostrado ser superior a otras funciones de distribución de probabilidad de precipitación diaria (Selker & Haith, 1990).

### Temperatura

De acuerdo con Richardson (1981), los datos de temperatura y radiación solar son más fáciles de modelar estadísticamente que la precipitación, pues hay una menor proporción de observaciones con valores de cero y la distribución de estas variables es mucho menos sesgada que los datos de precipitación. La técnica utilizada en *ClimGen* para la generación de temperatura máxima y mínima es similar a la de *WGEN*, y se basa en el supuesto de que la temperatura es un proceso débilmente estacionario (Matalas, 1967).

Este enfoque considera que la temperatura máxima y mínima debe ser un continuo proceso estocástico multivariado con medias diarias y desviaciones estándar condicionadas por la precipitación (húmedo o seco) del día (Richardson, 1981). Series temporales de los elementos residuales de la temperatura máxima y mínima se obtienen a través de la eliminación de las medias periódicas y el ajuste de las desviaciones estándar; los elementos son analizados por la dependencia en el tiempo y una correlación cruzada (Castellvi & Söckle, 2001; Stöckle & Nelson, 2003; Stöckle et al., 2001).

### Calibración y validación de ClimGen

Por lo general, las estaciones climáticas de México concentradas en el *ERIC III* comprenden las variables temperatura máxima y mínima, precipitación y evaporación con periodos considerables de tiempo, mas no así con otras variables como velocidad de viento, radiación solar y humedad relativa, entre otras. Por ello, la validación y calibración de *ClimGen* se realizó en específico para temperaturas y precipitación (cuadro 1). Esta herramienta computacional es de libre acceso y se puede obtener con previo registro en la siguiente dirección electrónica: [http://sites.bsyse.wsu.edu/cs\\_suite/ClimGen/documentation/demos.html](http://sites.bsyse.wsu.edu/cs_suite/ClimGen/documentation/demos.html).

La interfaz de *ClimGen* para la generación de datos climáticos se engloba en dos principales procedimientos: 1) el sitio de análisis y 2) la generación de los datos. En el procedimiento uno, primero se elige un directorio de salida al cual serán dirigidos los resultados; hecho esto, el siguiente paso es parametrizar el sitio de análisis. Esto comprende ingresar los siguientes datos: a) ubicación geográfica del lugar, es decir, se introducen las coordenadas de localización (latitud-longitud, grados decimales o UTM) y altitud; b) la región de ubicación con respecto al nivel medio del mar (*ClimGen* trae consigo una base de datos mundial de la cual se puede elegir el país de forma direccionada), y c) una breve descripción de la estación meteorológica.

En este mismo apartado se deben ingresar de forma manual los datos históricos año por año hasta completar el registro histórico de información; este proceso computacional es el que demanda más tiempo, al ingresar la información de forma reiterada. Una ventaja es que al ingresar un año con registros (completos o incompletos) se pueden observar de forma tabular y de forma gráfica (línea de tiempo), por lo que se puede hacer un segundo análisis gráfico exploratorio de la información. Integrados todos los registros históricos, *ClimGen* realiza un análisis estadístico (media, desviación estándar, valor máximo, mínimo, suma), para determinar si es posible producir la generación de los datos

o no; para ejecutar estimaciones de temperatura y precipitación es requerido un mínimo de 10 y 25 años, respectivamente.

Realizado lo anterior, el segundo procedimiento es la generación de los datos. Para ello se elige el archivo generado en el procedimiento anterior y el periodo a generar. En este contexto, *ClimGen* ofrece dos opciones de simulación: a) generar los datos faltantes, conservando los datos históricos y estimando sólo los valores ausentes; b) reemplazar totalmente todos los registros. *ClimGen* permite exportar los datos de salida en diferentes formatos, dependiendo el interés del usuario (*CropSyst*/*SWAT*, *LARS-WG*, entre otros). El formato *LARS-WG* fue considerado en el presente trabajo, por generar los datos en forma tabular, lo que permite una mejor manipulación de la información.

La validación del modelo consistió en la parametrización de la estación meteorológica Tepehuanes (figura 1). Se identificaron los años con la mayor y la menor precipitación histórica (1923, con 756 mm, y 1929, con 242.6 mm) y, de modo aleatorio, se eliminaron registros, con el fin de evaluar la capacidad predictiva de datos faltantes en la serie. De las dos opciones de simulación que ofrece *ClimGen*: a) generar los datos faltantes, conservando los datos históricos y estimando sólo los valores ausentes, y b) reemplazar totalmente todos los registros; se emplearon las dos opciones. Se hizo la simulación correspondiente y se graficaron sólo los datos observados *versus* los calculados; asimismo, se obtuvieron los parámetros estadísticos de los datos observados con respecto a los calculados y los generados empleando el programa estadístico *SPSS v. 17.0*, para validar la eficiencia de estimación de datos climáticos en un periodo de un año (1923, considerado el año con mayor precipitación, y 1929, año con el registro histórico de menor precipitación).

De manera consecutiva se hizo la parametrización para un periodo de tiempo más extenso; se consideraron 66 años de datos observados para la estación Tepehuanes (10084) y 45 años para la estación El Tarahumar (10026). Stöckle

et al. (1999) sugieren un mínimo de 20 años de datos meteorológicos históricos para completar la parametrización de variables climatológicas a ser producidos por un generador climático, y de preferencia entre 30 y 45 años. Por tal razón, los registros disponibles cumplieron con lo requerido para la estimación de datos faltantes de temperatura y precipitación en las dos estaciones.

Las estaciones validadas anteriormente comprendieron una condición de precipitación media (676 mm). Para evaluar la eficiencia del modelo en condiciones medioambientales contrastantes se analizaron las estaciones Riito, en el estado de Sonora (precipitación promedio anual de 52.5 mm), y Francisco Rueda del estado de Tabasco (2 414 mm de precipitación promedio anual) (IMTA, 2009). En las estaciones analizadas, los periodos con mayor ausencia de información fueron de tres meses, con variaciones interanuales de un mes sin registros. Por último, se hicieron gráficos de dispersión y se obtuvieron los parámetros estadísticos de los valores históricos observados y los calculados por *ClimGen*.

## Resultados

### Estimación de datos faltantes de un año

Existe una marcada similitud en la tendencia entre los datos observados y calculados para dos años con condiciones climáticas diferentes, un año con el mayor registro de precipitación (figura 2a) y otro con el menor registro (figura 2b), ambos correspondientes a la estación climatológica Tepehuanes, en el estado de Durango.

Los parámetros estadísticos de los datos observados con respecto a los calculados y generados por *ClimGen* para el año 1923 determinan que la varianza de los datos observados es mantenida en los datos calculados por el generador climático, no así para los datos generados (cuadro 2).

Analizando el año 1929, que históricamente registró la menor precipitación la varianza

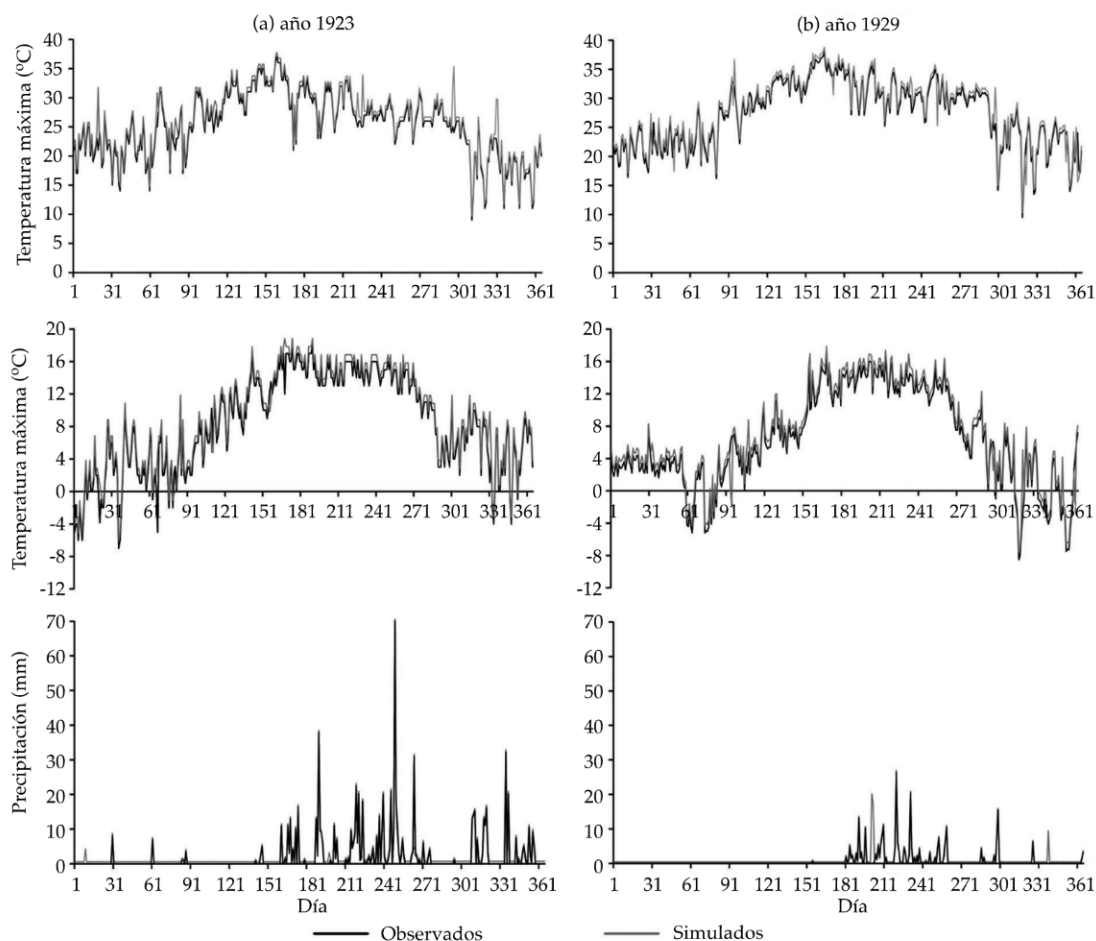


Figura 2. Series anuales de temperatura y precipitación observada y simulada empleando ClimGen para la estación Tepehuanes.

de los datos observados, es mantenida en los datos calculados; sin embargo, para los datos generados, esta varianza disminuye (cuadro 3).

### Estimación de datos faltantes mayores a un año

Al considerar periodos más extensos de información para el cálculo de datos faltantes, estación Tepehuanes (66 años) y El Tarahumar (45 años), bajo el criterio de  $r^2$ , la variable temperatura máxima presentó  $r^2=0.98$ , temperatura

mínima  $r^2=0.90$  y precipitación  $r^2=0.96$  (figura 3). Respecto a la estación El Tarahumar, la temperatura máxima mostró una  $r^2=0.98$ , temperatura mínima  $r^2=0.99$  y precipitación de  $r^2=0.93$  (figura 3). Al conservar los registros históricos y calcular sólo los valores faltantes, la varianza de los datos observados es mantenida en los datos calculados por el generador climático (cuadro 4).

Al analizar dos estaciones con condiciones mediomambientales contrastantes a la estación Tepehuanes y El Tarahumar; Riito en el estado de



Cuadro 2. Parámetros estadísticos de los datos observados *versus* calculados y generados del año 1923.

	N	Media	Varianza	Desviación estándar	Máximo	Mínimo	R	R <sup>2</sup>
Tmáx_o	365	25.26	29.65	5.44	37	9		
Tmáx_c	365	25.45	30.36	5.51	37	9	0.97	0.95
Tmáx_g	365	26.71	33.91	5.82	44	12	0.52	0.27
Tmín_o	365	8.38	33.87	5.82	18	-7		
Tmín_c	365	8.36	34.31	5.85	18	-7	0.98	0.97
Tmín_g	365	7.27	42.40	6.51	20	-7	0.78	0.61
PP_o	365	2.1	37.55	6.12	70	0		
PP_c	365	2.12	37.53	6.12	70	0	0.99	0.98
PP_g	365	0.99	12.23	3.49	42	0	0.07	0.005

<sup>1</sup>La denotación "o" significa datos observados.<sup>2</sup>La denotación "s" significa datos calculados por ClimGen.<sup>3</sup>La denotación "g" significa datos generados por ClimGen.Cuadro 3. Parámetros estadísticos de los datos observados *versus* calculados y generados del año 1929.

	N	Media	Varianza	Desviación estándar	Máximo	Mínimo	R	R <sup>2</sup>
Tmáx_o	365	25.26	29.65	5.44	37	9		
Tmáx_c	365	25.45	30.36	5.51	37	9	0.97	0.95
Tmáx_g	365	26.71	33.91	5.82	44	12	0.54	0.30
Tmín_o	365	8.38	33.87	5.82	18	-7		
Tmín_c	365	8.36	34.31	5.86	18	-7	0.98	0.97
Tmín_g	365	7.27	42.40	6.51	20	-7	0.71	0.50
PP_o	365	2.10	37.55	6.13	70	0		
PP_c	365	2.12	37.53	6.13	70	0	0.86	0.75
PP_g	365	0.99	12.23	3.50	42	0	0.16	0.026

<sup>1</sup>La denotación "o" significa datos observados.<sup>2</sup>La denotación "s" significa datos calculados por ClimGen.<sup>3</sup>La denotación "g" significa datos generados por ClimGen.

Sonora y Francisco Rueda en el estado de Tabasco, los gráficos de dispersión muestran que la variable temperatura máxima presentó  $r^2=0.96$ , temperatura mínima  $r^2=0.95$  y precipitación  $r^2=0.99$ , con base en la estación Riito. Respecto a la estación Francisco Rueda, la temperatura máxima mostró una  $r^2=0.96$ , temperatura mínima  $r^2=0.98$  y precipitación de  $r^2=0.97$  (figura 4). Tanto para las estaciones climatológicas de El Tarahumar y Tepehuanes como para las estaciones Riito y Francisco Rueda, si se conservan los registros históricos en la base de datos y sólo

se generan los datos faltantes, la variancia de los datos observados es mantenida en los datos calculados por el generador climático (cuadros 4 y 5).

## Discusión

El empleo de *ClimGen* se ha extendido con éxito debido a sus múltiples aplicaciones para generar series de tiempo, como un paso previo para la aplicación de otros modelos (p. ej., *CRO-PWAT*) (Sarangi & Kumar, 2006; Bal, Choudhury,



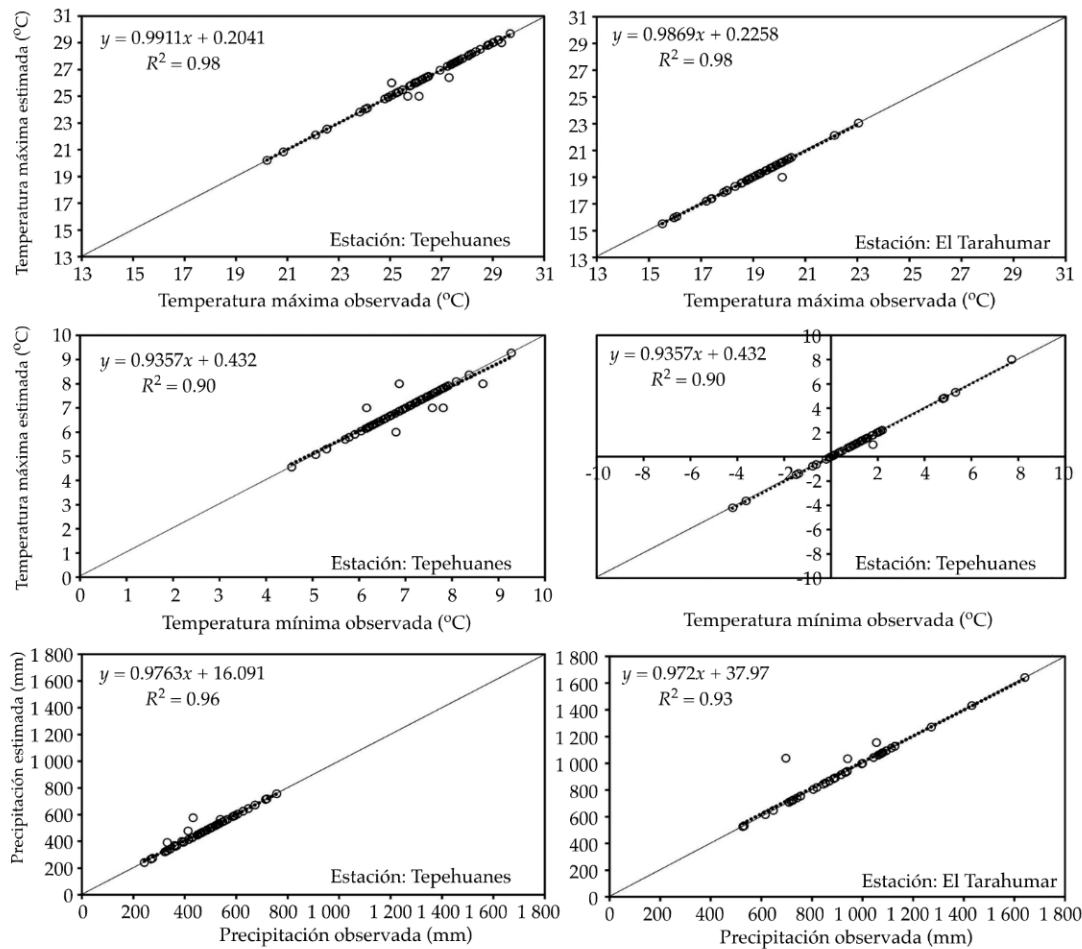


Figura 3. Variables climáticas (Temperatura máxima, mínima y precipitación) observadas y estimadas por ClimGen bajo el criterio de  $r^2$ .

Cuadro 4. Estadística descriptiva de los datos observados *versus* calculados por ClimGen para la estación climatológica El Tarahumar y Tepehuanes.

	Tepehuanes					
	Temperatura máxima		Temperatura mínima		Precipitación	
	Observada	Calculada	Observada	Calculada	Observada	Calculada
Media	26.5	26.4	6.9	6.9	472	477
Desviación estándar	1.95	1.94	0.83	0.82	110.86	110.23
Varianza	3.79	3.78	0.69	0.67	12 289.74	12 151.19
	El Tarahumar					
	Temperatura máxima		Temperatura mínima		Precipitación	
	Observada	Calculada	Observada	Calculada	Observada	Calculada
	Observada	Calculada	Observada	Calculada	Observada	Calculada
Media	19.1	19.1	1.0	1.0	914	926
Desviación estándar	1.43	1.43	2.02	2.04	218.09	218.83
Varianza	2.06	4.08	47 565.33	2.03	4.16	47 888.62

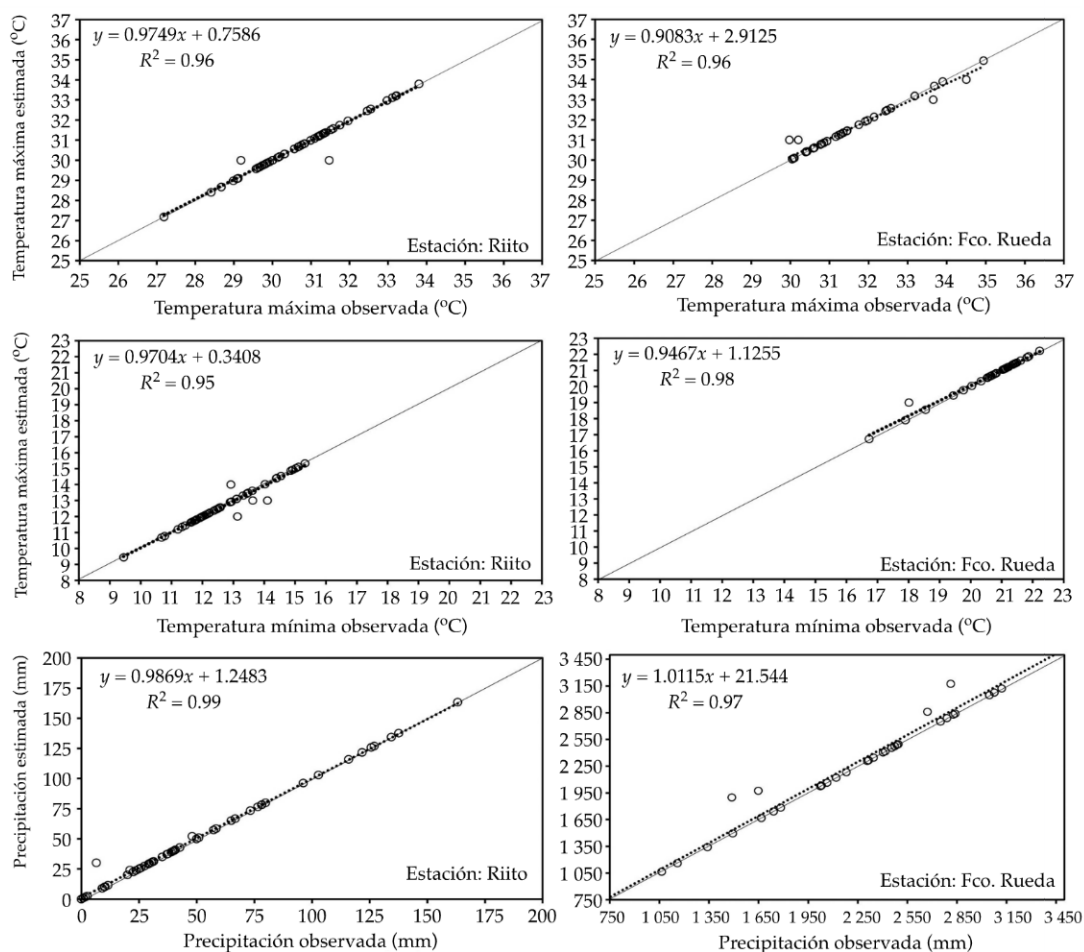


Figura 4. Temperatura máxima, mínima, y precipitación observada y estimada por ClimGen bajo el criterio de  $r^2$ .

Cuadro 5. Estadística descriptiva de los datos observados *versus* calculados por ClimGen para la estación climatológica Riito y Francisco Rueda.

	Riito					
	Temperatura máxima		Temperatura mínima		Precipitación	
	Observada	Calculada	Observada	Calculada	Observada	Calculada
Media	31.2	31.2	13.0	12.9	49	50
Desviación estándar	1.73	1.67	1.40	1.40	39.63	39.28
Varianza	3.00	2.78	1.97	1.95	1 570.50	1 543.12
	Francisco Rueda					
	Temperatura máxima		Temperatura mínima		Precipitación	
	Observada	Calculada	Observada	Calculada	Observada	Calculada
	Observada	Calculada	Observada	Calculada	Observada	Calculada
Media	31.6	31.6	20.6	20.6	2 189.2	2 235.8
Desviación estándar	1.32	1.22	1.17	1.12	755.24	773.44
Varianza	1.74	1.49	1.38	1.26	570 388.82	598 213.57

Sood, Jalota, & Singh, 2008; Tingem et al., 2007; McKague, Rudra & Ogilvie, 2003), para estimar variables climáticas (Sarangi, Madramootoo, & Koundal, 2008; Carbajal et al., 2010; Safeq & Fares, 2011), pero muy pocos trabajos se han enfocado en evaluar la eficiencia del modelo para la estimación de datos faltantes en una serie de datos (Moradi, Nosrati, & Eslamian, 2007; Carbajal et al., 2010).

Moradi et al. (2007), en Irán, evaluaron cinco estaciones climáticas para estimar temperatura máxima y mínima. Los coeficientes de correlación indicaron que los datos observados y los estimados presentaron un buen ajuste ( $r = 0.82$ ,  $r = 0.72$ ,  $r = 0.79$  y  $r = 0.78$ , excepto para la temperatura mínima de una estación,  $r = 0.10$ ). Sin embargo, recomienda su uso en la estimación de datos faltantes en regiones con presencia de media a baja precipitación.

Safeq y Fares (2011) evaluaron el modelo para cuatro estaciones bajo condiciones tropicales, encontrando que su eficiencia de estimación variaba según el sitio y la variable climática (p. ej., precipitación, velocidad de viento, temperatura). De los cuatro sitios, *ClimGen* sólo fue capaz de reproducir la precipitación observada para una estación y en dos lugares subestimó esta variable. Por el contrario, estimó de manera razonable la temperatura máxima, mínima y la velocidad de viento.

En el caso de la estación climatológica Tepihuanes, los análisis realizados para un periodo de un año muestran que *ClimGen* es capaz de producir estimaciones estadísticamente significativas para las variables temperatura máxima ( $r^2 = 0.95$ ), temperatura mínima ( $r^2 = 0.97$ ) y precipitación ( $r^2 = 0.98$ ), considerando el año con la mayor presencia de precipitación (1923), y  $r^2 = 0.97$  para la variable temperatura máxima,  $r^2 = 0.95$  para la temperatura mínima y  $r^2 = 0.75$  para la precipitación para el año 1929, que históricamente registró la menor precipitación. Dichas estimaciones se obtienen al conservar los datos históricos en la serie de datos; por el contrario, al no conservar los datos observados, la variabilidad conservada

disminuye para todas las variables de manera importante, temperatura máxima  $r^2 = 0.27$ , temperatura mínima  $r^2 = 0.61$  y precipitación  $r^2 = 0.005$  para el año 1923. Asimismo, para el año 1929, la temperatura máxima presenta una  $r^2 = 0.30$ , temperatura mínima  $r^2 = 0.50$ , y precipitación  $r^2 = 0.026$ .

Aunque se analizaron dos años con la mayor y menor precipitación histórica, el modelo estimó de forma aceptable las dos condiciones medioambientales estudiadas y conserva la varianza de los datos observados en los datos calculados. Sin embargo, al no considerar los datos existentes y generar por completo los 365 días, la eficiencia de estimación disminuyó. Esto corrobora los resultados encontrados por Moradi et al. (2007), y Safeq y Fares (2011), en los cuales se reportan estimaciones de buena calidad para ciertos parámetros climáticos y errores sustanciales para algunos sitios de estudio.

Considerando periodos mayores a un año, los gráficos de dispersión y el valor correspondiente de  $r^2$  indican una buena eficiencia predictiva para la variable temperatura máxima ( $r^2 = 0.98$ ) en ambas estaciones, para temperatura mínima  $r^2 = 0.90$  en la estación Tepihuanes, y  $r^2 = 0.99$  para El Tarahumar; para la precipitación una  $r^2 = 0.96$  y  $r^2 = 0.93$ , respectivamente. Estos valores muestran que para regiones de condición media de precipitación, el generador climático estimó valores de buena calidad que representan el clima de la región.

Por otro lado, en la parte norte del país, la estación meteorológica Ritto del estado de Sonora, con precipitación de 52.5 mm promedio anual, que representa las condiciones de aridez del país, las estimaciones para la variable temperatura máxima mostraron una  $r^2 = 0.96$ , temperatura mínima  $r^2 = 0.95$  y precipitación  $r^2 = 0.99$ . Asimismo, la estación Francisco Rueda del estado de Tabasco, con una precipitación de 2 414 mm promedio anual, la cual representa la condición húmeda, muestra una buena eficiencia de estimación para la variable temperatura máxima ( $r^2 = 0.96$ ). En ambas estaciones, para temperatura mínima  $r^2 = 0.98$  y

$r^2 = 0.97$  para la precipitación, respectivamente. Estos resultados demuestran la viabilidad del generador climático para la estimación de datos faltantes con independencia de las condiciones medioambientales de las estaciones de estudio. Sin embargo, se debe tener presente que las estimaciones se atribuyen principalmente a dos factores: a) cantidad de información histórica en la base de datos, y b) mayor o menor presencia de datos faltantes. El periodo máximo de datos faltantes fue de tres meses consecutivos o variaciones interanuales de un mes sin registros, los cuales se encontraban en la parte inicial, media y final del año. No se evaluó el generar un año o más sin información, ya que no fueron los planteamientos perseguidos en este estudio, aunque se debe tener presente que acorde con los resultados, cuando se generó un año completo, la varianza disminuyó.

Sin embargo, son resultados considerablemente aceptables en relación con otros trabajos, en los cuales se valoró la eficiencia del generador climático. Por ejemplo, en el altiplano peruano, se evaluó el modelo *ClimGen* para la estimación de datos faltantes de precipitación, obteniendo valores de  $r^2 = 0.70$  y  $r^2 = 0.71$  en dos estaciones, y de  $r^2 = 0.40$  para la tercera estación (Carbajal et al., 2010). Safeq y Fares (2011) demostraron que el desempeño del generador varía incluso entre estaciones de una misma región geográfica y para una variable climática en particular, reportando estimaciones confiables para diferentes variables climáticas (p. ej., temperatura máxima, temperatura mínima y velocidad de viento).

La fortaleza de contar con una serie extensa y completa de registros históricos de clima permite parametrizar modelos en los cuales es indispensable que no se tengan datos faltantes en la serie de datos (Esquivel et al., 2013a; Esquivel, Bueno, Sánchez, Velásquez, & Delgado, 2013b) o que ayuden a calibrar estudios en los cuales las estaciones climatológicas desempeñan un papel fundamental (Cerano, Villaueva, Cervantes, Trucios, & Guerrero, 2013). Las comparaciones entre los datos observados y calculados son una parte esencial de la validación del modelo, pues

proporcionan información valiosa sobre el comportamiento de los datos y determina en dónde es capaz de funcionar o no de forma adecuada. *ClimGen* demostró en general funcionar bien en las estaciones de estudio y se observó que mayor cantidad de datos observados y menor cantidad de datos faltantes en una base de datos fortalece la parametrización del modelo e incrementa la calidad de las estimación de datos.

## Conclusiones

*ClimGen* mostró un buen desempeño al calcular datos faltantes de la serie de tiempo histórica, lo cual indica que los datos estimados por el generador climático son representativos de los datos históricos del clima. Al calcular datos faltantes, es importante conservar los datos históricos observados de la estación y sólo calcular los datos faltantes; esto permite conservar la varianza de los datos calculados (temperatura máxima;  $r^2 = 0.95$ ) versus si se generan todos los datos por completo (temperatura máxima;  $r^2 = 0.27$ ). Este tipo de procedimientos permite extender los registros históricos observados de una estación en particular, lo que ampliará la posibilidad de utilizar estaciones climatológicas para su uso en estudios hidrológicos, ecológicos, paleoclimáticos y de cultivo, entre otros. Sin embargo, es necesario remarcar que se debe tener cuidado al utilizar los datos generados por el modelo y analizarlos gráficamente para identificar datos que sobreestimen o subestimen de manera significativa los registros históricos, antes de que sean utilizados como datos de entrada en una aplicación en particular. También parte importante del algoritmo mostrado en este estudio es la validación previa de la información climática a utilizar, dado que de esta calidad dependerá la información generada. El estudio también indica que es posible utilizar registros anuales de variables climáticas con información faltante, haciendo de lado la práctica común de desechar estos datos, mermando la amplitud de la información disponible para fines de modelación de procesos.



## Referencias

- Bal, S. K., Choudhury, B. U., Sood, A., Jalota, S. K., & Singh, H. (2008). Evaluation of climgen model to generate weather parameters under different climatic situations in Punjab. *Journal of Agrometeorology*, 1(1), 39-45.
- Carbajal, C. M., Yarlequé, C., Posadas, A., Silvestre, E., Mejía, A., & Quiroz, R. (2010). Reconstrucción de datos faltantes de precipitación pluvial diaria mediante la Transformada Wavelet. *Revista peruana geo-atmósferica RPGA*, 2, 76-88.
- Castellvi, F., & Stöckle, C. O. (2001). Comparing the performance of WGEN and ClimGen in the generation of temperature and solar radiation. *Transactions of American Society of Agricultural and Biological Engineers (ASABE)*, 44, 1683-1687.
- Cerano, P. J., Villaveva, D. J., Cervantes, M. R., Trucios, C. R., & Guerrero, S. J. L. (2013). Reconstrucción de sequías fuertes en el parque nacional "Pico de Tancitaro, Michoacán". *Revista Chapingo Serie Zonas Áridas*, 12(2), 57-62.
- Conagua (2012). *Climatología estadística*. Comisión Nacional del Agua. Agosto 28 de 2012. Citado el ocho de enero de 2014. Recuperado de <http://smn.cna.gob.mx/climatologia/Estadistica/10084.pdf>; <http://smn.cna.gob.mx/climatologia/Estadistica/10026.pdf>; <http://smn.cna.gob.mx/climatologia/Estadistica/26076.pdf>; <http://smn.cna.gob.mx/climatologia/Estadistica/27015.pdf>.
- Danuso, F., & Della, M. V. (1997). *CLIMAK Reference manual* (pp. 148). DPVTA. Udine, Italy: University of Udine, Italy.
- Esquivel, A. G., Sánchez, C. I., Velásquez, V. M. A., López, S. A., López, L. R., & Bueno, H. P. (2013a). Modelación del escurrimiento en una subcuenca del trópico húmedo de México y su análisis mediante índices de eficiencia predictiva. *Agrofaz*, 13(2), 113-118.
- Esquivel, A. G., Bueno, H. P., Sánchez, C. I., Velásquez, V. M. A., & Delgado, R. G. (2013b). Eficiencia predictiva de modelos hidrológicos para cuencas poco instrumentadas. *Memoria III Congreso Nacional de Manejo de Cuencas Hidrográficas*. Morelia, Michoacán México.
- Flanagan, D. C., & Livingston, S. J. (1995). *WEPP User Summary*. NSERL Report No. 11. West Lafayette, USA: USDA-ARS National Soil Erosion Research Laboratory.
- Hoogenboom, G. (2000). Contribution of Agro-Meteorology to the Simulation of Crop Production and its Applications. *Agricultural and Forest Meteorology*, 103, 137-157.
- Hoogenboom, G., & Soltani, A. (2003). Minimum Data Requirements for Parameter Estimation of Stochastic Weather Generators. *Climate Research*, 25, 109-119.
- IMTA (2009). *Extractor Rápido de Información Climatológica v. 3.0 (ERIC III). Base de datos y software*. Jiutepec, México: Instituto Mexicano de Tecnología del Agua.
- Johnson, G. L., Hanson, C. L., Hardegree, S. P., & Ballard, E. B. (1996). Stochastic Weather Simulation: Overview and Analysis of Two Commonly Used Models. *Journal of Applied Meteorology and Climatology*, 35, 1878-1896.
- Matalas, N. C. (1967). Mathematical Assessment of Synthetic Hydrology. *Water Resources Research*, 3(4), 937-945.
- Mckague, K., Rudra, R., & Ogilvie, J. ClimGen - A Convenient Weather Generation Tool for Canadian Climate Stations. Paper No. 03-118. *Proc. CSAE/SCGR 2003 Meeting*, 6-9 July, Montreal, Québec.
- Moradi, I., Nosrati, K., & Eslamian, S. (2007). Evaluation of the RadEst and Climgen Stochastic Weather Generators for Low-Medium Rainfall Regions. *Journal of Applied Sciences*, 7(19), 2900-2903.
- Nicks, A. D., Richardson, C. W., & Williams, J. R. (1990). Evaluation of the EPIC Model Weather Generator (pp. 105-124). In A. N. Sharpley & J. R. Williams (Eds.). *EPIC - Erosion/Productivity Impact Calculator, 1. Model Documentation*, USDA Technical Bulletin 1768. Washington, DC: Government Printing Office.
- Richardson, C. W. (1981). Stochastic Simulation of Daily Precipitation, Temperature, and Solar Radiation. *Water Resources Research*, 17(1), 182-190.
- Richardson, C. W., & Wright, D. A. (1984). *WGEN: A Model for Generating Weather Variables*. US Department of Agriculture, Agricultural Research Service, ARS-8.
- Sarangi, A., & Kumar, A. (2006). *Trend Analysis of long Term Rainfall Data for Crop Planning as an Adaptation Measure to Climate Change* (pp. 1-6). 60<sup>th</sup> International Executive Council Meeting & 5<sup>th</sup> Asian Regional Conference, New Delhi, India.
- Sarangi, A., Madramootoo, C. A., & Koundal, K. R. (2008). *Generation of Regional Water Harvesting Potential Scenarios using ClimGen Model* (pp. 139-142). International Symposium Water Harvesting, Brining Green Revolution to Rainfed Areas, New Delhi, India.
- Safeeq, M., & Fares, A. (2011). Accuracy Evaluation of ClimGen Weather Generator and Daily to Hourly Disaggregation Methods in Tropical Conditions. *Theoretical and Applied Climatology*, 106, 321-341.
- Selker, J. S., & Haith, D. A. (1990). Development and Testing of Single Parameter Precipitation Distribution. *Water Resources Research*, 26(11), 2733-2740.
- Stöckle, C. O., & Nelson, R. (2003). *Cropping System Simulation Model user's Manual* (p. 235). Washington, DC: Washing State University.
- Stöckle, C. O., Nelson, R., Donatelli, M., & Castellvi, F. (2001). *ClimGen: A Flexible Weather Generation Program*. 2nd International Symposium Modelling Cropping Systems, Florence, Italy, 16-18 July.
- Stöckle, C. O., Campbell, G. S., & Nelson, R. (1999). *ClimGen Manual* (28 pp.). Pullman, USA: Biological Systems Engineering Department, Washington State University.
- Tingem, M., Rivington, M., Azam-Ali, S. N., & Colls, J. J. (2008). Climate Variability and Maize Production in Cameroon: Simulating the Effects of Extreme Dry and Wet Years. *Singapore Journal of Tropical Geography*, 29(3), 357-370.

- Tingem, M., Rivington, M., Azam-Ali, S. & Colls, J. (2007). Assessment of the Climgen stochastic weather generator at Cameroon sites. *African Journal of Environmental Science and Technology*, 1(4), 86-92.
- Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics – Transactions ASME*, 18(3), 293-297.

## Dirección institucional de los autores

M.C. Gerardo Esquivel Arriaga

Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias  
km 6.5 margen derecha canal Sacramento  
35140 Gómez Palacio, Durango, MÉXICO  
Teléfono: +52 (871) 1590 104  
esquivel.gerardo@inifap.gob.mx

M.C. Julián Cerano Paredes

Universidad Nacional Autónoma de México  
Instituto de Geografía  
Departamento de Geografía Física  
Circuito de la Investigación Científica, Ciudad Universitaria, Coyoacán  
04510 México, D.F., MÉXICO  
Teléfono: +52 (871) 1590 104  
cerano.julian@inifap.gob.mx

Dr. Ignacio Sánchez Cohen

Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias  
km 6.5 margen derecha canal Sacramento  
35140 Gómez Palacio, Durango, MÉXICO  
Teléfono: +52 (871) 1590 104  
sanchez.ignacio@inifap.gob.mx

Dr. Armando López Santos

Universidad Autónoma Chapingo  
Unidad Regional Universitaria de Zonas Áridas  
Dom. Conocido s/n Bermejillo  
35230 Durango, MÉXICO  
Teléfono: +52 (872) 7760 190, 60  
armando.lopezsantos@gmail.com

Dr. Oscar Gerardo Gutiérrez Ruacho

Universidad Estatal de Sonora  
Rosales núm. 189. Col. Centro  
83001 Hermosillo, Sonora, MÉXICO  
Teléfono: +52 (662) 2153 778  
ruachogr@hotmail.com  
oscar.gutierrez@ues.mx



Haga clic aquí para escribir al autor