

DOI: 10.24850/j-tyca-14-01-10

Notas

Calidad de datos de estaciones meteorológicas automáticas ubicadas en tres estados de México

Data quality of automatic meteorological stations located in three states of Mexico

Rocío Cervantes-Osomio¹, ORCID: <https://orcid.org/0000-0003-2597-4517>

Ramón Arteaga-Ramírez², ORCID: <https://orcid.org/0000-0001-9459-3588>

Mario Alberto Vázquez-Peña³, ORCID: <https://orcid.org/0000-0003-2084-7420>

Waldo Ojeda-Bustamante⁴, ORCID: <https://orcid.org/0000-0001-7183-9637>

¹Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Cevamex, Texcoco, Estado de México, México, rcervanteso@hotmail.com

²Universidad Autónoma Chapingo, Departamento de Irrigación, Estado de México, México, arteagarr@gmail.com



³Universidad Autónoma Chapingo, Departamento de Irrigación, Estado de México, México, mvazquezp@gmail.com

⁴Instituto Mexicano de Tecnología del Agua, Jiutepec, Morelos, México, w.ojeda@riego.mx

Autora para correspondencia: Rocío Cervantes-Osornio, rcervanteso@hotmail.com

Resumen

El objetivo del presente trabajo fue aplicar algunos procedimientos para evaluar la calidad de datos provenientes de diferentes estaciones meteorológicas automáticas (EMA) ubicadas en tres estados de la república mexicana. Se obtuvieron datos meteorológicos diarios de estaciones automáticas ubicadas en los siguientes puntos: una en el estado de México, cinco en Guanajuato, y una en Sinaloa. Los procedimientos utilizados para la temperatura máxima, media y mínima, y radiación global fueron la prueba de rango flexible o efeméride meteorológica; para las variables temperatura y humedad relativa se empleó la prueba de consistencia interna; a la variable temperatura se le aplicó la prueba de salto térmico y de ceros; a la humedad relativa, la prueba de consistencia interna y la de persistencia. Para la estación ubicada en Sinaloa, la variable radiación global diaria, en la prueba de rango flexible, presentó 23.25 % de datos erróneos. Los resultados de los



procedimientos aplicados a datos históricos de las EMA apoyan en determinar la confiabilidad de la información meteorológica de éstas para su uso.

Palabras clave: calidad, datos meteorológicos, humedad relativa, temperatura, radiación solar.

Abstract

The objective of the present work was to apply some procedures to evaluate the quality of data coming from different automatic meteorological stations (EMA) located in three states of the Mexican Republic. Daily meteorological data were obtained from automatic stations: one in the State of Mexico, five in Guanajuato, and one in Sinaloa. The procedures that were used were: for the maximum, average, and minimum temperature and global radiation, the flexible range test or meteorological event was applied; for the temperature and relative humidity variables, the internal consistency test was used; to the variable temperature, the thermal jump and the zeros test were applied; to the relative humidity, the internal consistency, and the persistence test. For the station located in Sinaloa, for the variable daily global radiation, on the flexible range test, it presented 23.25 % of erroneous data. The results of the procedures applied to historical data of the EMAs support determining the reliability of the meteorological information of these for use.



Keywords: Quality, meteorological data, relative humidity, temperature, solar radiation.

Recibido: 11/08/2020

Aceptado: 18/09/2021

Introducción

Se considera que los datos meteorológicos son de calidad cuando responden a determinadas normas tanto explícitas como implícitas (WMO, 1996), que dependen de las diferentes aplicaciones donde se va a utilizar la información (Estévez, 2008), algunas de las cuales pueden distinguir anomalías de temperatura y precipitación bajo escenarios de cambio climático (Campos-Aranda, 2020), o descubrir patrones en datos climatológicos, particularmente en la temperatura (Bulege-Gutiérrez & Custodio, 2020). Se aplican distintos métodos a los datos meteorológicos para asegurar su calidad: mantenimiento periódico y correcto de las estaciones y de los sensores en campo, la calibración de estos últimos y la validación de los datos mediante procedimientos estadísticos (Estévez,



Gavilán, & Girálvez, 2011). Obrien y Keefer (1985) propusieron un conjunto de tres reglas que se basan en la informática, las cuales fueron aplicadas por Meek y Hatfiel (1994) para validar los datos meteorológicos, y que incluyen el cálculo de límites altos, bajos, fijos o dinámicos para cada variable; y el uso de una tasa fija o dinámica de los límites de cambio para cada variable e identificación de un cambio continuo, no observado, en un límite de tiempo. Estévez, García-Marín, Morábito y Cavagnaro (2016) llevaron a cabo un estudio en Mendoza, provincia de Argentina, y aplicaron diversos métodos (detección y marcado de datos vacíos; prueba de rango o límite fijo y dinámico; prueba de salto; de consistencia interna y persistencia) para validar datos provenientes de las variables: radiación solar diaria; temperatura máxima, mínima y media diaria; humedad relativa máxima, mínima y media diaria; velocidad del viento diaria máxima, mínima y media.

Shafer, Fiebrich y Arndt (2000) aplicaron para la temperatura (a 1.5 y 9 m) y la humedad relativa máxima las pruebas de rango, paso, persistencia, espacial y comparación de instrumentos similares, a datos registrados, algunos cada cinco minutos, y otros a 15 y 30 minutos, de 115 estaciones de Oklahoma; personal del *staff* encontró durante el periodo 1998 más de 50 “ráfagas de calor”, eventos relativamente comunes, las cuales afectaron una o dos estaciones, y aparecen como condiciones de calor y sequedad anómalas durante la noche; además, registraron datos con problemas que se almacenaron de manera automática, pero el juicio humano fue capaz de discernir sobre estas

indicaciones y preservó los registros de tales eventos para su estudio futuro.

El objetivo del presente trabajo fue determinar la calidad de mediciones meteorológicas con diferentes procedimientos para estaciones automáticas de tres diferentes localidades (Guanajuato, Sinaloa y Estado de México).

Materiales y métodos

Los datos meteorológicos diarios utilizados son los siguientes: radiación global (RG); temperatura máxima (T_{max}), media (T_{med}) y mínima (T_{min}); humedad relativa máxima (HR_{max}), media (HR_{med}) y mínima (HR_{min}). La latitud, longitud, altitud y periodo de estudio de las estaciones automáticas utilizadas se presentan en la Tabla 1.



Tabla 1. Coordenadas geográficas y periodo de observación de las estaciones de estudio.

Estado	Estación	Latitud	Longitud	Altitud, msnm	Periodo de estudio
Estado de México	Chapingo	19° 29' 05"	98° 53' 11"	2 250	23/1/2004-31/12/2013
Guanajuato	Agroeduca	21° 00' 39.2"	101° 39' 4.9"	1 777	1/1/2001-31/12/2006
Guanajuato	El Fuerte	20° 37' 26.4"	101° 12' 28.3"	1 716	1/1/2001-31/12/2006
Guanajuato	El Cardonal	21° 11' 39.7"	100° 46' 39"	1 912	1/1/2001-31/12/2006
Guanajuato	El Galomo	21° 04' 2.7"	100° 27' 14.3"	2 048	1/1/2001-31/12/2006
Guanajuato	El Jardín	21° 15' 52.5"	100° 35' 30.1"	1 968	1/1/2001-31/12/2006
Sinaloa	Santa Rosa 1	25° 45' 03"	108° 57' 21"	40	23/4/1997-15/5/2001

Para la validación de los datos meteorológicos fue necesario hacer un análisis previo de la base de datos, lo que significó detectar espacios vacíos (blancos). Para evaluar los datos diarios de temperatura máxima (T_{max}), media (T_{med}), mínima (T_{min}); radiación global (RG); y humedad relativa máxima (HR_{max}), media (HR_{med}) y mínima (HR_{min}) se utilizó la prueba de rango rígido. Para la temperatura máxima, media y mínima y

radiación global se empleó la prueba de rango flexible o efeméride meteorológica. Para la humedad relativa no se encontraron datos extremos en las normales climatológicas, debido a esto fue irrealizable la prueba de rango flexible. A las variables de temperatura (T_{max} , T_{med} y T_{min}) y humedad relativa (HR_{max} , HR_{med} , HR_{min}) se les aplicó la prueba de consistencia interna; a la temperatura, la prueba de salto térmico ($T_{max}(d) - T_{min}(d)$) y prueba de ceros; a la humedad relativa (HR_{max} , HR_{med} y HR_{min}) y radiación global, la prueba de persistencia. Para aplicar la prueba de rango flexible o efemérides se necesita contar con valores extremos normales y debido a que las estaciones automáticas utilizadas en el estudio tienen datos de 10 años o menos, se seleccionaron de las normales climatológicas (NC) —publicadas por el Servicio Meteorológico Nacional (SMN)— estaciones cercanas, con características similares en cuanto a altitud, latitud y longitud (Tabla 2), de las cuales se tomaron los valores necesarios para esta prueba; sólo la estación Chapingo está en las normales climatológicas.

Tabla 2. Estaciones seleccionadas de las normales climatológicas (NC).

Estado	Estación	Estación de las NC	Latitud	Longitud	Altitud, msnm
Guanajuato	Agroeduca	San Francisco Durán	21° 00' 19"	101° 37' 16"	1790
Guanajuato	El Fuerte	Salamanca (DGE)	20° 33' 50"	101° 11' 54"	1723
Guanajuato	El Cardonal	Peñuelitas	21° 06' 16"	100° 52' 33"	1915
Guanajuato	El Galomo	San José Iturbide (DGE)	20° 59' 52"	100° 23' 07"	2099
Guanajuato	El Jardín	Lourdes	21° 17' 24"	100° 42' 07"	1995
Sinaloa	Santa Rosa 1	El Zopilote	25° 43' 50"	108° 20' 45"	37

Prueba de blancos. Se detectan los espacios vacíos o datos faltantes para las diferentes variables en las bases de datos de las estaciones que se trabajaron.

Prueba de rango rígido y rango flexible (efemérides meteorológicas) (Estévez *et al.*, 2011). Se utilizaron los límites físicos propuestos por Feng, Hu y Qian (2004) y UNE 500540 (2004) debido a que no se tuvo acceso a los datos de las especificaciones de los sensores respectivos de las tres localidades de estudio. Además, esta prueba incorpora la verificación de efemérides meteorológicas (límite dinámico o rango flexible), donde cada valor se compara con el valor extremo registrado históricamente de determinado lugar.

Prueba de rango flexible para radiación solar. Se utilizó la radiación solar diaria (R_s o RG) de las tres localidades, los cuales se representaron gráficamente, con relación a la radiación que se espera que ocurra bajo condiciones de cielos despejados o limpios (R_{so}), y se calcula (Allen, 2008) con:

$$R_{so} K_t \cdot R_a \quad (1)$$

Donde K_t es el índice de "claridad" y R_a es la radiación extraterrestre con las mismas unidades de la R_s ($\text{MJ m}^{-2} \text{d}^{-1}$ o $\text{MJ m}^{-2} \text{h}^{-1}$ o W m^{-2}). De acuerdo con Allen (2008) y de la metodología de la ASCE-EWRI (2005), se tiene lo siguiente: $K_t = K_B + K_D$, donde K_B es el índice de claridad del haz directo de la radiación (adimensional) y K_D es el índice de transmisividad para la radiación difusa, adimensional; entonces:

$$R_{so} = (K_B + K_D) \cdot R_a \quad (2)$$

La estimación de K_B propuesta por Majumdar, Mathur y Kaushik (1972), Allen (1996), y Allen, Smith, Pereira y Perrier (1998) es aplicada por ASCE-EWRI (2005), con la mejora de los coeficientes desarrollados de la evaluación del ASCE Task Committee para datos de R_s de cerca de 50 sitios meteorológicos en EUA; para K_B se tiene:

$$K_B = 0.98 \exp \left[\frac{-0.00146P}{K_t \sin \beta} - 0.075 \left(\frac{W}{\sin \beta} \right)^{0.4} \right] \quad (3)$$

Donde:

K_t es el coeficiente de turbidez (adimensional) $-0 < K_t \leq 1.0-$, donde $K_t = 1.0$ para aire limpio, y $K_t \leq 0.5$ para el aire turbio, con polvo o contaminado.

P es la presión atmosférica $-(\text{kPa})-$ se estima con la elevación del sitio de acuerdo con Allen *et al.* (1998).

β es el ángulo del sol sobre el horizonte (rad).

W es el agua precipitable en la atmósfera (mm).

El valor para K_t varía en función de la época del año y la limpieza de la atmósfera debido a la precipitación. En general, para la predicción de la radiación neta (R_n) y envolventes R_{so} se recomienda $K_t = 1.0$. El $\sin \beta$ en la Ecuación (3) debe limitarse a valores ≥ 0.01 para la estabilidad computacional.

El agua precipitable se estima como:

$$W = 0.14e_a P + 2.1 \quad (4)$$

Donde e_a es la presión de vapor actual del aire (a 2 m) (kPa).



El índice de transmisividad para radiación difusa se estima a partir de K_B (según Allen, 1996):

$$K_D = 0.35 - 0.36K_B \text{ para } K_B \geq 0.15$$

$$K_D = 0.18 + 0.82K_B \text{ para } K_B < 0.15$$

Para periodos de tiempo diarios (24 horas), el valor medio de β se calcula de acuerdo con Allen (1996), como:

$$(\sin \beta)_{24} = \sin \left[0.85 + 0.3\varphi \sin \left(\frac{2\pi}{365} J - 1.39 \right) - 0.42\varphi^2 \right] \quad (5)$$

Donde $(\sin \beta)_{24}$ es el promedio durante el periodo de luz diurna (rad); φ , la latitud (rad), y J es el día juliano (adimensional).

La variable $(\sin \beta)_{24}$ se usa en lugar de $\sin \beta$ en la Ecuación (3) y representa el promedio ponderado del ángulo del sol durante las horas de luz diurna. El valor para $(\sin \beta)_{24}$ debe ser ≥ 0 .

Prueba de consistencia interna (Reek, Doty, & Owen, 1992). Asegura la coherencia meteorológica de la variable en cuestión, además verifica que la temperatura máxima (valor instantáneo) sea mayor que la

temperatura media (valor promedio), y ésta mayor que la temperatura mínima (valor instantáneo) (Ecuación (8), Tabla 3); lo mismo sucede para la variable humedad relativa máxima, mínima y media (Ecuación (14), Tabla 3). Se trata de comprobar la coherencia de las tres variables que se almacenan diariamente (Estévez *et al.*, 2011).

Tabla 3. Ecuaciones y pruebas aplicadas en las diferentes estaciones a las variables meteorológicas en las localidades de estudio.

Variable	Prueba de	Ecuación	Núm. ec.	
Temperatura máxima, mínima y media	Rango rígido	$-35 < T_{max}, T_{min}, T_{med} < 55$	(6)	
	Rango flexible (efeméride meteorológica)	$T_B < T_{max}, T_{min}, T_{med} < T_A$	(7)	
	Consistencia interna		$T_{max} > T_{med} > T_{min}$	(8)
			$T_{max}(d) > T_{min}(d-1)$	(9)
			$T_{min}(d) \leq T_{max}(d-1)$	(10)
	Salto térmico diario	$dT(d) < 23.8\text{ }^{\circ}\text{C}$	(11)	
Ceros	$T_{min}(d) = T_{min}(d-1) = T_{min}(d-2) = 0$	(12)		
Humedad relativa máxima, mínima y media	Rango rígido	$0.8 < HR_{max}, HR_{min}, HR_{med} < 100$	(13)	
	Consistencia interna	$HR_{max} > HR_{med} > HR_{min}$	(14)	
	Persistencia		$HR_{max}(d) \neq HR_{max}(d-1) \neq HR_{max}(d-2)$	(15)
			$HR_{min}(d) \neq HR_{min}(d-1) \neq HR_{min}(d-2)$	(16)
			$HR_{med}(d) \neq HR_{med}(d-1) \neq HR_{med}(d-2)$	(17)

Radiación solar	De persistencia	$RG(d) \neq RG(d-1) \neq RG(d-2)$	(18)
	Rango flexible o efeméride meteorológica	$RG < 1.1 \cdot R_{so}$	(19)

T_{max} es la temperatura máxima ($^{\circ}\text{C}$); T_{med} es la temperatura media ($^{\circ}\text{C}$); T_{min} es la temperatura mínima ($^{\circ}\text{C}$); HR_{max} es la humedad relativa máxima (%); HR_{med} es la humedad relativa media (%); HR_{min} es la humedad relativa mínima (%); (d) es el día evaluado para la variable en cuestión (T_{max} , T_{med} , T_{min} , HR_{max} , HR_{med} , HR_{min}); ($d-1$) es el día anterior evaluado para la variable en cuestión (adimensional); ($d-2$) es la evaluación hace dos días para la variable en cuestión (adimensional); $dT(d) = T_{max}(d) - T_{min}(d)$ es el salto térmico diario ($^{\circ}\text{C}$); T_A es el valor de temperatura extrema más alta registrada en el SMN para la estación en cuestión ($^{\circ}\text{C}$); T_B es el valor de temperatura mínima más baja registrada en el SMN para la estación en cuestión ($^{\circ}\text{C}$); RG es la radiación global ($\text{MJ m}^{-2}\text{día}^{-1}$); R_{so} es la radiación global diaria bajo condiciones de cielo claro ($\text{MJ m}^{-2}\text{día}^{-1}$).

Fuente: Elaborado de Estévez *et al.*, 2011; AENOR, 2004; Reek *et al.*, 1992

Las ecuaciones (9) y (10) (Tabla 3) verifican que la temperatura máxima sea mayor que la mínima del día anterior, y que la temperatura mínima sea menor o igual que la máxima del día anterior.

Prueba de salto térmico diario. Verifica que la diferencia de T_{max} y T_{min} para un día cualquiera no supere un umbral determinado (Araya-López, 2011; Reek *et al.*, 1992).

Prueba de persistencia. Verifica si existen más de dos registros consecutivos con el mismo valor de humedad relativa, máxima, media, mínima o radiación global para los datos diarios (Meek & Hatfield, 1994).

Prueba de ceros. Tiene la finalidad de verificar registros continuos iguales a cero de la variable temperatura mínima (Schroeder, Burgett, Haynie, & Sonmez, 2005).

Resultados y discusión

Prueba de blancos. Para las tres localidades se detectaron el número de registros en blanco (campos vacíos) existentes en cada caso; se agruparon por año y variable. Para las cinco estaciones en Guanajuato y para Santa Rosa I AC, en Sinaloa, no se presentaron campos vacíos (0 %) de todas las variables y para los años estudiados.

Para la estación Chapingo, en el Estado de México, se registraron 160 espacios vacíos para cada conjunto de 3 631 datos para esta estación; cinco años no se presentan por datos perdidos; el mayor porcentaje de campos vacíos se encuentra en 2012. Cabe destacar que el número de registros en blanco es el mismo para cada variable

meteorológica (RG , T_{max} , T_{med} , T_{min} , HR_{max} , HR_{med} y HR_{min}), lo cual indica que son problemas en el equipo de la EMA.

Prueba de rango rígido. Las estaciones de Guanajuato, Estado de México (Chapingo), y Sinaloa (Santa Rosa I AC) para las variables T_{max} , T_{min} y T_{med} no determinaron datos fuera del rango de la Ecuación (6).

Para las variables humedad relativa máxima y media, los resultados muestran que no hay dato que esté fuera del rango de la Ecuación (13) para las cinco estaciones de Guanajuato; sin embargo, para la humedad relativa mínima, las estaciones El Galomo y El Jardín presentan un dato sospechoso cada una: 22/05/2006 (0.05 %) y 07/09/2001 (0.05 %), respectivamente. Para Chapingo, las variables HR_{max} , HR_{med} y HR_{min} , todos sus datos del periodo de estudio, exceptuando los blancos, cumplen con la condición de la Ecuación (13), y para Santa Rosa I AC para la HR_{min} existe un dato (06/02/1999) sospechoso (0.07 %).

Meek y Hatfield (1994) reportan que existen patrones cuando algún dato se está midiendo correctamente por el sensor, y que un dato inusual debido a un evento climático extremo se puede verificar, ya que existen patrones entre variables relacionadas, por ejemplo, en un dato de humedad relativa baja con uno reportado para ese mismo día de temperatura muy alta en el aire incluso se presentan estos patrones en estaciones relativamente cercanas. Meek y Hatfield (1994) reportan cinco días de datos perdidos debido a que el sensor al parecer se congeló en intervalos diversos en ese periodo; comentan que existen factores de

relación que muestran esto: 1) la temperatura del aire se reporta bajo 0° ; (2) la presión de vapor estuvo cerca del punto de saturación; (3) los datos de dirección del viento muestran variación, y (4) hay presencia de hielo en el trípode e instrumentos; comentan que en la práctica, estos datos no deben ser borrados sino únicamente marcados (*flagged*).

Prueba de rango flexible o efemérides. Con la Ecuación (7) se encontró que para Guanajuato, la variable T_{max} , y para las estaciones Agroeduca, El Cardonal, El Fuerte, El Galomo y El Jardín, todos los datos cumplen con las condiciones de esta ecuación, y se representan en la Figura 1 y Figura 2.

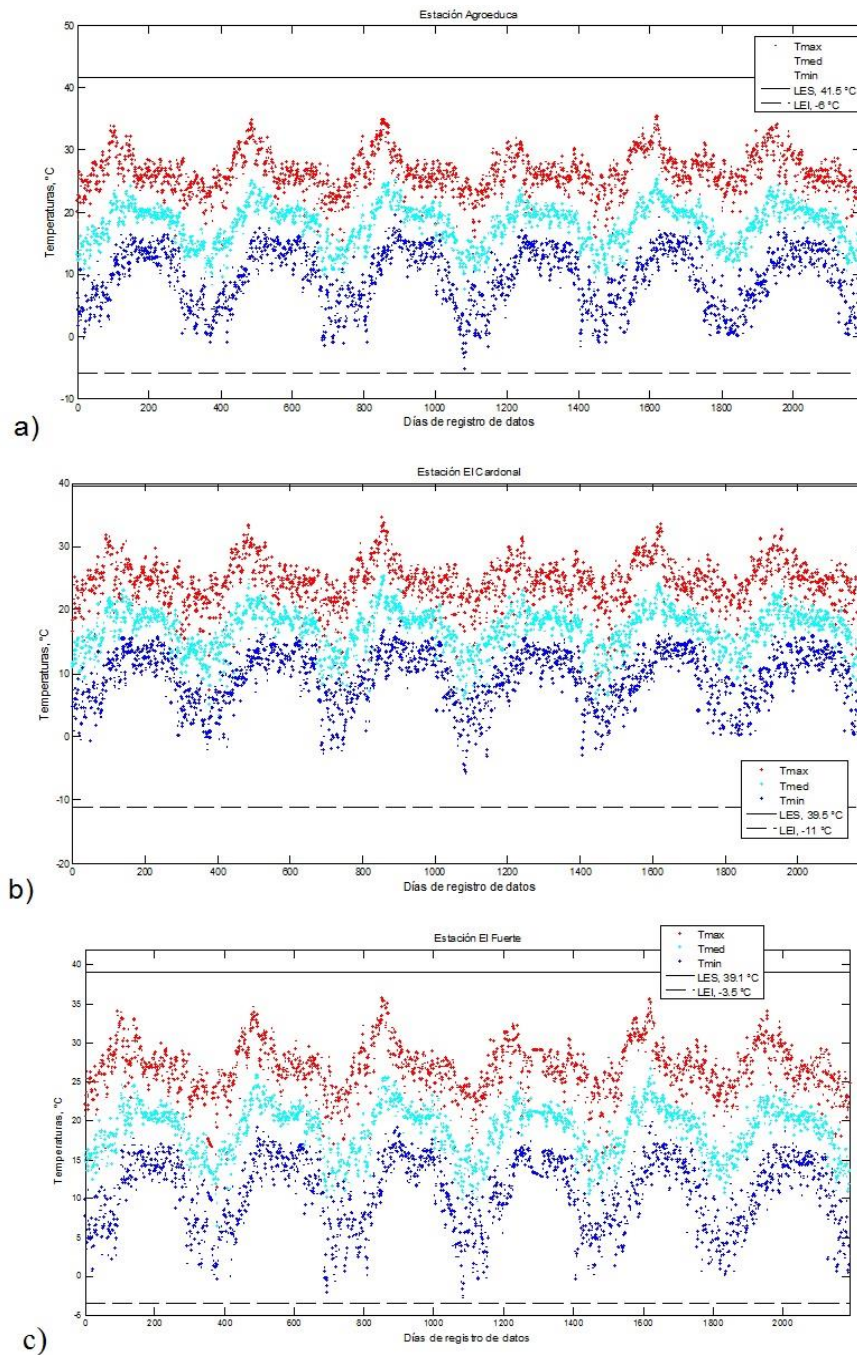


Figura 1. Variación de la T_{max} , T_{med} , T_{min} , localización de los valores



extremos para las estaciones a) Agroeduca ($-6 > T_{max,min,med} > 41.5$); b) El Cardonal ($-11 > T_{max,min,med} > 39.5$), y c) El Fuerte ($-3.5 > T_{max,min,med} > 39.1$). LES: límite extremo superior; LEI: límite extremo inferior.

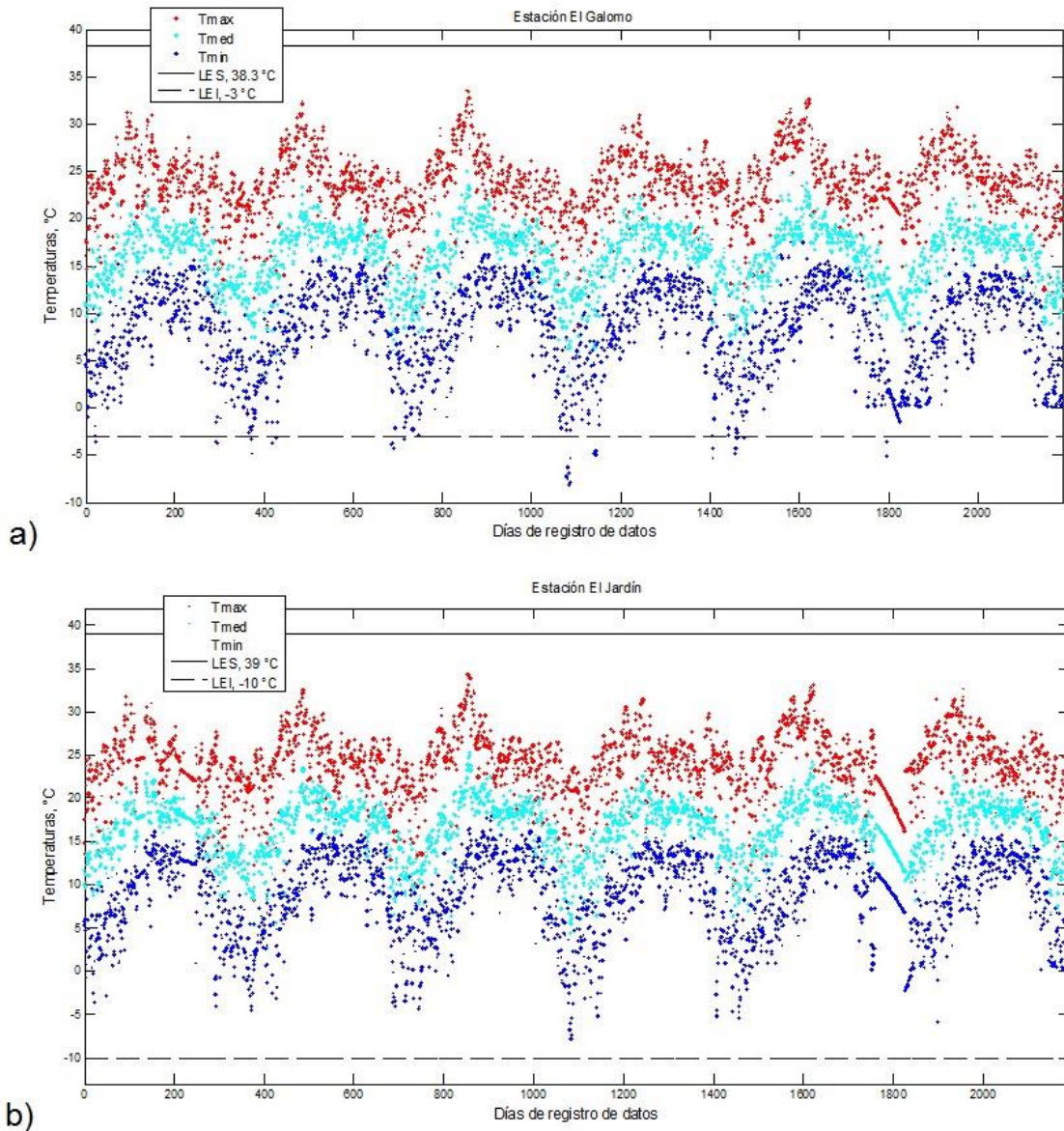


Figura 2. Variación de la T_{max} , T_{med} , T_{min} , localización de los valores extremos para las estaciones a) El Galomo ($-3 > T_{max,min,med} > 38.3$) y b) El Jardín ($-10 > T_{max,min,med} > 39$). LES: límite extremo superior; LEI: límite extremo inferior.

Para la variable de temperatura mínima (T_{min}) para las estaciones Agroeduca, El Cardonal, el Fuerte y el Jardín, todos los datos cumplen con las condiciones de valores extremos de las normales climatológicas por estación; sin embargo, para la estación El Galomo hay 34 datos (1.55 %) que no cumplen con la condición $-3 > T_{max,min,med} > 38.3$ (Ecuación (7)), colocándolos en categoría de sospechosos. Para Chapingo, para T_{max} y T_{med} , todos los datos cumplen con la condición $-9 < T_{max,min,med} < 38$ (Figura 3a); sin embargo, para T_{min} hay un dato espurio: 08/08/2011 (0.03 %). Para Santa Rosa I AC, sólo la variable T_{med} no presentó datos etiquetados como sospechosos, pero los datos de las variables T_{max} y T_{min} presentaron en cada conjunto un dato sospechoso con las fechas 14/10/1997 (0.07 %) y 6/02/1999 (0.07 %), respectivamente, que no cumplieron con la condición $0 < T_{max,min,med} < 43$ (Figura 3b). Meek y Hatfiel (1994) aplicaron la regla LIM ($Ta_{min}(d) \leq Ta(d) \leq Ta_{max}(d)$), que es el equivalente a la Ecuación (7) en este trabajo. Para el territorio de Walnut Creek reportaron un dato atípico, pero comentan que es un dato válido, al analizar el reporte de un evento climático extremo (lluvia tormentosa severa) donde toda el área experimentó una fuerte tormenta de nieve.

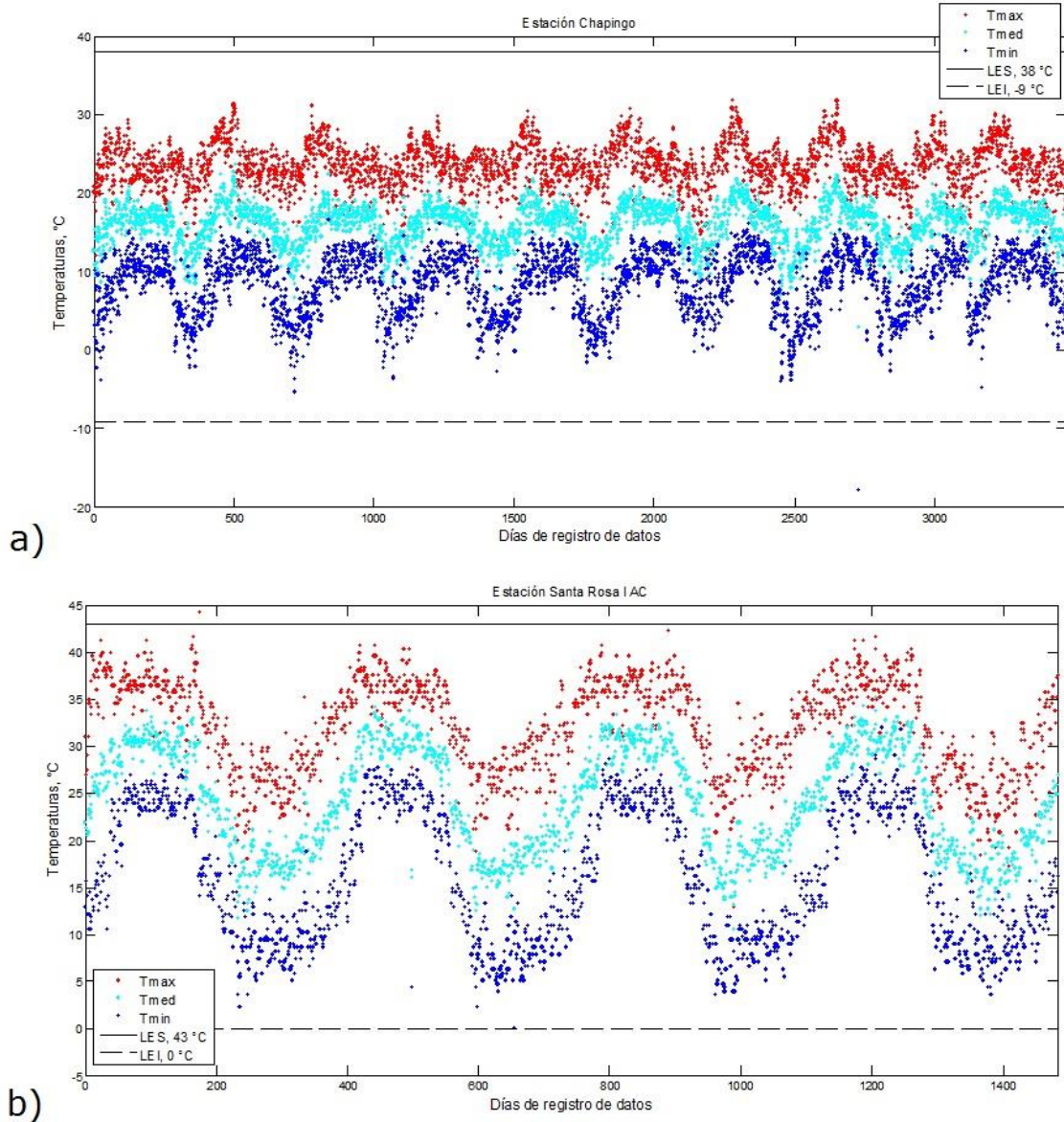


Figura 3. Variación de la T_{max} , T_{med} , T_{min} ; localización de los valores extremos para las estaciones a) Chapingo ($-9 < T_{max,min,med} < 38$) y b) Santa Rosa I AC ($0 < T_{max,min,med} < 43$). LES: límite extremo superior; LEI: límite extremo inferior.

De igual forma, Feng *et al.* (2004) reportan errores de 0.000035 % para T_{max} para la prueba similar a la Ecuación (7) que utilizan estos autores $-89.4 < Td < 93.9$ °F, y de 0.000009 % para T_{min} , donde utilizaron los límites extremos de temperatura de Kubecka (2001) y Gleason (2002).

Prueba de rango flexible (efeméride) para radiación global.

Con respecto a la radiación global para esta prueba (Ecuación (19)), se obtuvieron los valores de R_{so} para las estaciones de las diferentes localidades. Las envolventes de R_{so} se muestran en la Figura 4, Figura 5 y Figura 6; aquí se observa la condición $RG < 1.1R_{so}$ que utilizan Estévez *et al.* (2011), Allen (1996) y Moradi (2007), y que considera como sospechosos los datos que están por encima de $1.1 \cdot R_{so}$.

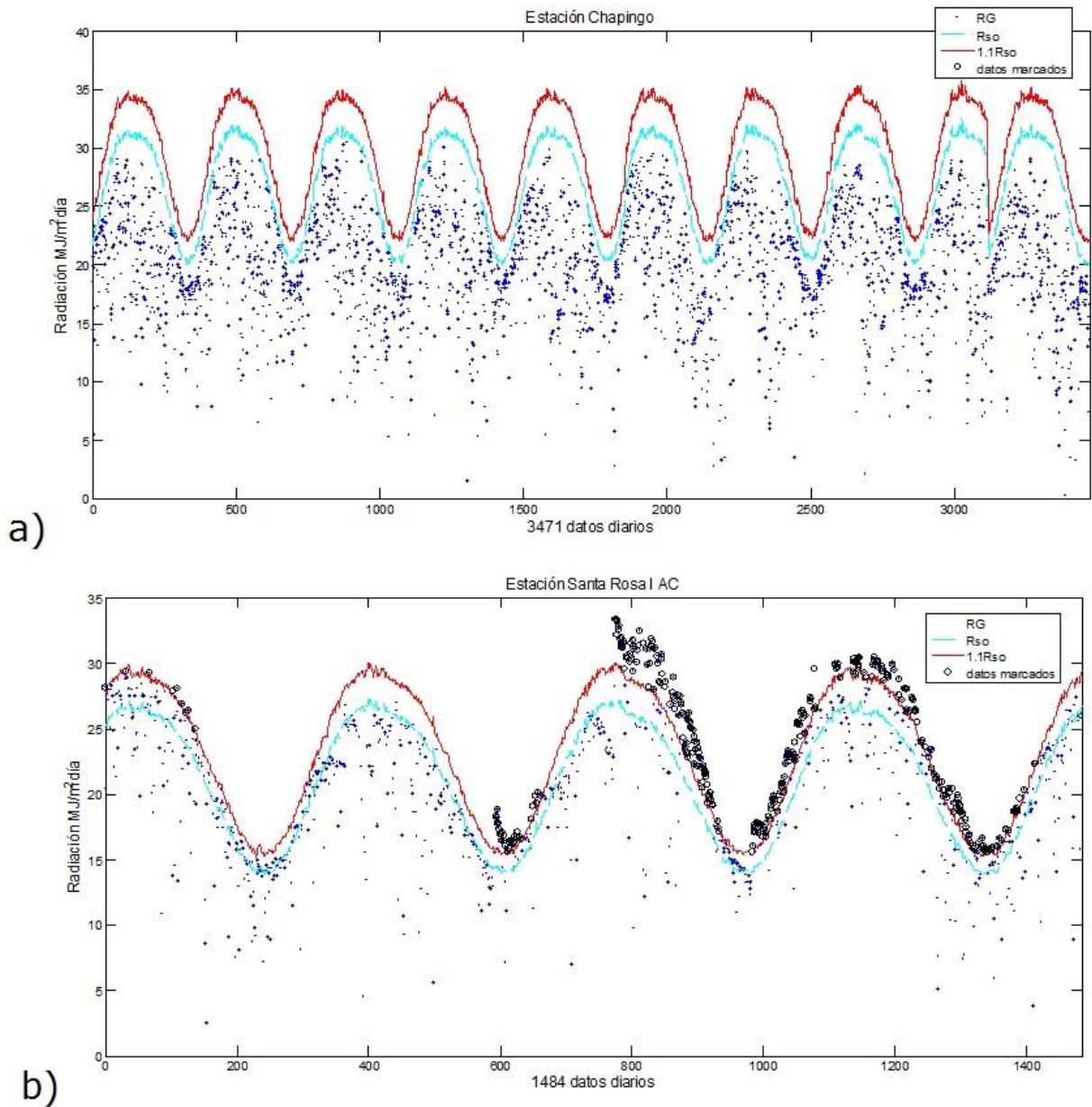


Figura 4. Evolventes R_{So} , $1.1 \cdot R_{So}$ y RG para las estaciones a) Chapingo y b) Santa Rosa I AC.

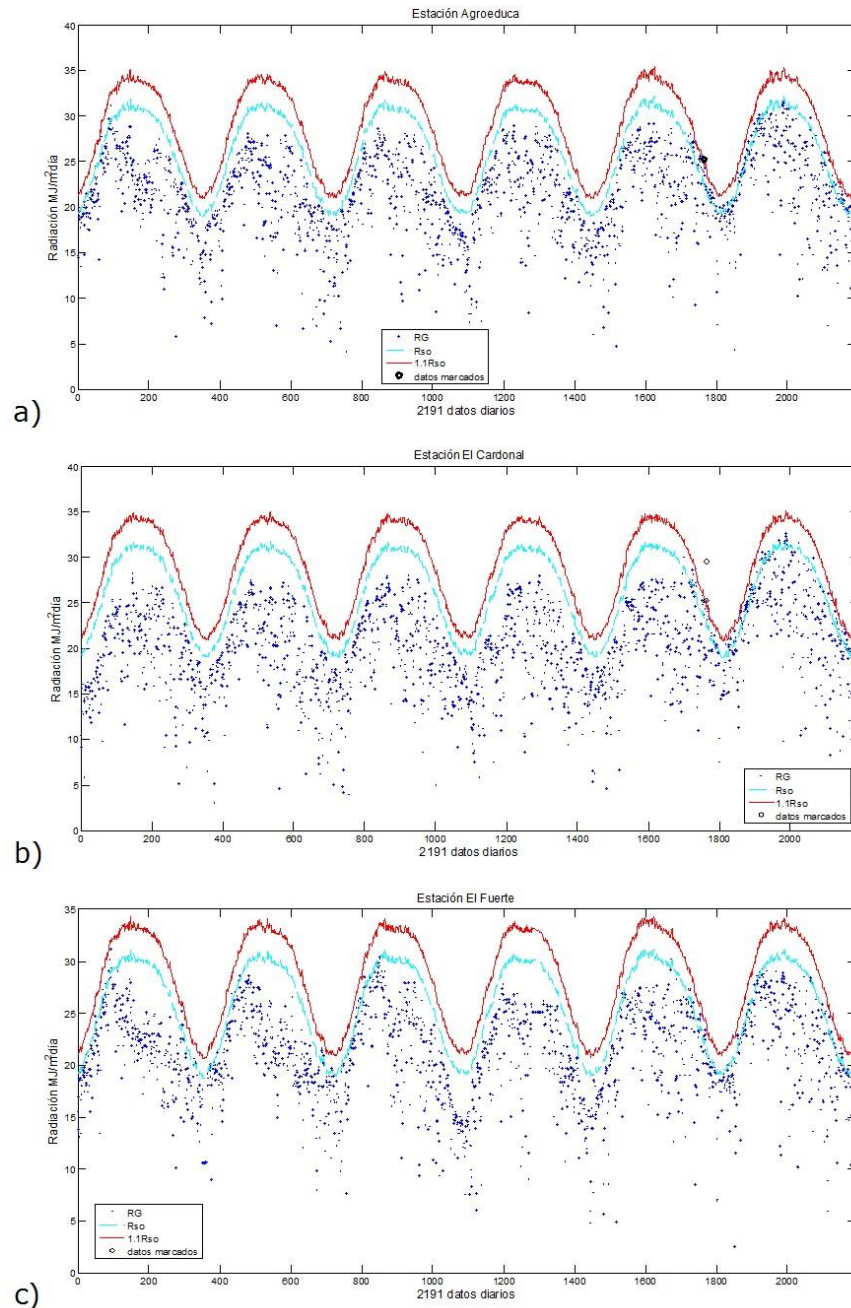


Figura 5. Evolventes R_{so} , $1.1 \cdot R_{so}$ y RG para las estaciones a) Agroeduca, b) El Cardonal y c) El Fuerte.

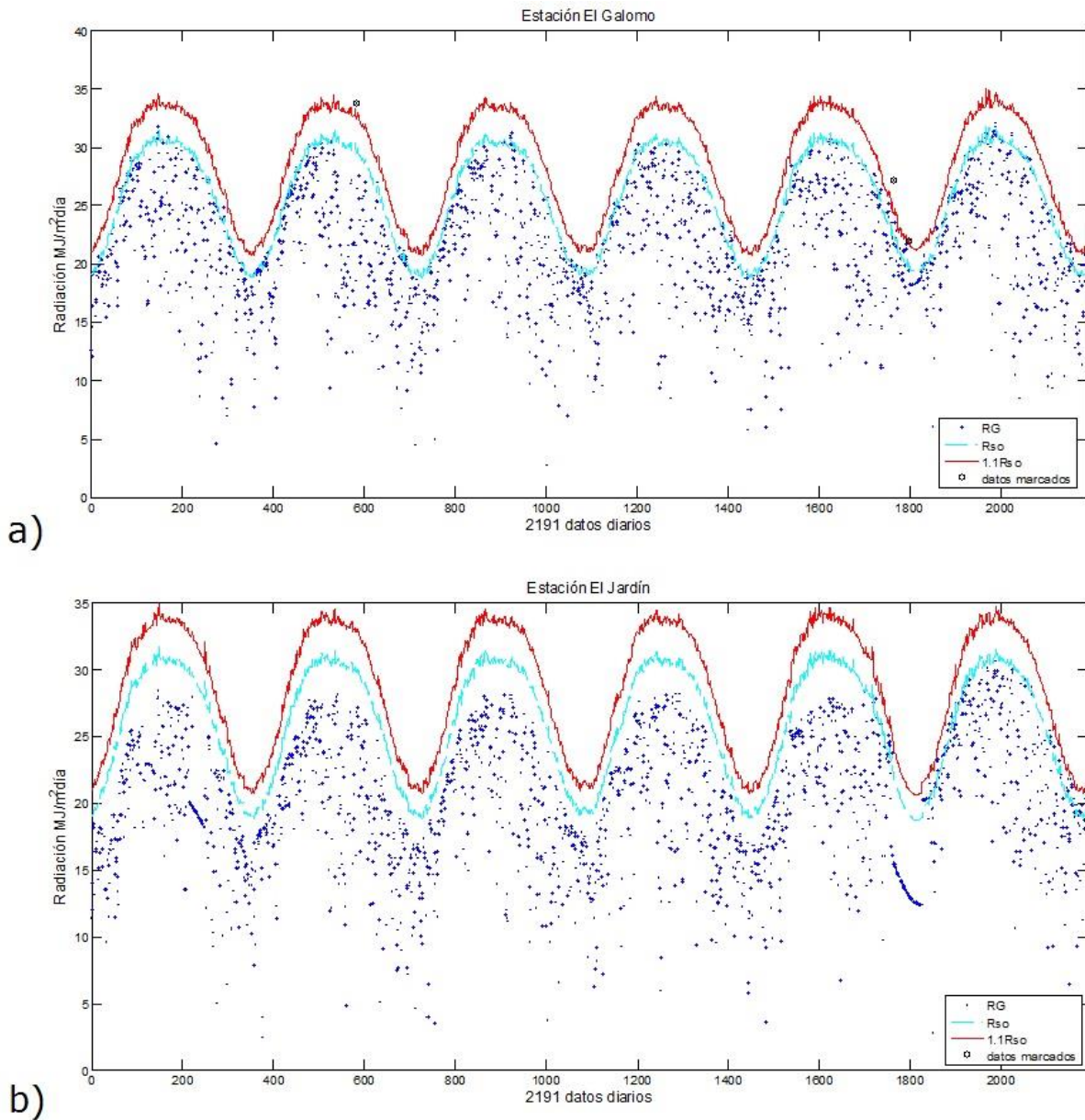


Figura 6. Evolventes R_{s0} , $1.1 \cdot R_{s0}$ y RG para las estaciones a) El Galomo y b) El Jardín.

Para Chapingo (Figura 4a), todos los datos cumplen con la condición de la Ecuación (19). Para Santa Rosa I AC (Figura 4b), 345 (23.25 %) datos son no confiables (datos marcados); en 1998, el sensor de esta estación empezó a presentar desajustes, específicamente en el dato 775 y en el año 1999, el daño en el sensor se presentó de manera permanente, ya que marcaba cantidades de radiación solar en las noches cuando debería haber indicado valores de cero. Estévez *et al.* (2011) encontraron, para la estación Las Cabezas de San Juan en Sevilla, múltiples mediciones en un 20 % mayor que la condición de la Ecuación (19) en el periodo de 2002 a 2006.

Con respecto a Guanajuato (Figura 5 y Figura 6), para las siguientes estaciones hay datos sospechosos: Agroeduca, un dato (0.05 %), 31/10/2005; El Cardonal, dos datos (0.09 %), 30/10/2005 y 31/10/2005; El Galomo, tres datos: (0.14 %), 7/8/2002, 31/10/2005 y 2/12/2005. Se observa que Agroeduca, El Cardonal y El Galomo reportan un dato atípico (31/10/2005), lo cual indica que se presentó esta condición en la zona en general; ello puede deberse a cielos muy despejados, o claros y limpios, por lo que no fue una medición errónea del sensor; así, estos datos sólo deben ser marcados, pero son válidos.

Prueba de consistencia interna. En Guanajuato, para la variable temperatura, en todas las estaciones se cumple con la condición de la Ecuación (8); de igual manera, todos los datos cumplen con la condición de la Ecuación (10). Sin embargo, para la Ecuación (9), en la Estación El

Galomo existe un dato que no cumple con la condición dada (10/01/2004), y lo coloca en la categoría de inconsistente (0.05 %); los datos para las demás estaciones sí cumplen. En Chapingo, para la variable temperatura, cumple todas las condiciones de las ecuaciones (8), (9) y (10). En cuanto a Santa Rosa I, para la Ecuación (8) existen cuatro datos inconsistentes (0.27 %) debido a que marcaban T_{max} igual a T_{med} y T_{min} ; también para la Ecuación (9) hay un dato inconsistente (0.07 %). Para la Ecuación (10), todos los datos cumplen con esta condición.

Para la variable humedad relativa, todos los datos de las estaciones de Guanajuato son consistentes con la Ecuación (14); lo mismo se presenta para Chapingo. Pero para Santa Rosa I AC, para la variable humedad relativa (Ecuación (14)), existen seis datos inconsistentes (0.40 %).

Prueba de salto térmico. En El Galomo, Agroeduca, El Cardonal, El Fuerte y El Jardín, los datos (y porcentajes) que no cumplen con la condición de la Ecuación (11) son los siguientes: 95 valores (4.34 %); 126 valores (5.75 %); 39 valores (1.78 %); 59 valores (2.69 %), y 133 valores (6.07 %), respectivamente; para Chapingo existen 43 datos (1.24 %), y para Santa Rosa I, 50 datos (3.37 %) que se etiquetan como sospechosos. Araya-López y Alfaro-Martínez (2008) aplicaron una prueba de salto térmico horario a datos de Manzanillo y Pavas en Costa Rica, encontrando que entre datos horarios de temperatura existió un valor atípico del salto (15 °C); al comparar este dato con el de una estación

cercana quedó claro que el campo de temperaturas de la región no experimentó variaciones, por lo que se trataba de un dato erróneo.

Prueba de persistencia. Para la variable humedad relativa máxima para la localidad de Guanajuato, todas las estaciones presentaron datos que no cumplen con la condición de la Ecuación (15). Asimismo, en Chapingo y Santa Rosa I AC, los datos sospechosos son los siguientes: Agroeduca, 246 datos (11.23 %); el Cardonal, 116 datos (5.29 %); el Fuerte, 62 datos (2.83 %); el Galomo, 972 datos (44.34 %); el Jardín, 40 datos (1.82 %); Chapingo, 603 datos (17.37 %), y Santa Rosa I AC, 1149 datos (77.43 %).

Para la variable humedad relativa mínima todas las estaciones del estudio presentaron datos sospechosos, que no cumplieron con la condición de la Ecuación (16): Agroeduca, 20 datos (0.91 %); el Cardonal, 22 datos (1.004 %); el Fuerte, 22 datos (1.004 %); el Galomo, 8 datos (0.36 %); el Jardín, 10 datos (0.46 %); Chapingo, 349 datos (10.05 %); Santa Rosa I AC, 192 datos (12.94 %).

Allen (2008) comenta que valores de humedad relativa más bajos del 30 % en zonas subhúmedas no son valores esperados.

Para la variable humedad relativa media, las estaciones que presentaron datos sospechosos fueron las siguientes: el Cardonal, 6 datos (0.27 %); el Fuerte, 4 datos (0.18 %); el Galomo, 4 datos (0.18 %); el Jardín, 6 datos (0.27 %); Chapingo, 255 datos (7.35 %), y Santa Rosa I AC, 141 datos (9.50 %). Agroeduca no presentó datos erróneos para esta variable. Santa Rosa I, en esta prueba de persistencia, presentó valores

del 100 % en días continuos (uno tras otro), $HR_{max}(d) = HR_{max}(d - 1) = HR_{max}(d - 2) = 100 \%$, con un 77.43 % de valores que no cumplen la condición (Ecuación (15)), pero se argumenta que Santa Rosa I es una área de riego, por lo que presenta mediciones de 100 % en diversos intervalos de cinco minutos durante un día, sobre todo al amanecer, cuando se presenta el punto de rocío semejante a la temperatura mínima por lo que la humedad alcanza valores de 100 % o cercanos a este valor, y además está cerca de la costa (Allen *et al.*, 1998). En esta estación, de los datos tomados cada cinco minutos en un día se obtuvo el valor máximo y el mínimo. Dicho lo anterior, se debe tomar con reserva la denominación de datos "erróneos" para este porcentaje y específicamente para esta prueba; esto lo corroboran Meek y Hatfield (1994) cuando comentan que además de los eventos abruptos o severos, las observaciones marcadas (*flagged*) son poco frecuentes de manera diaria. Comentan también que existe una única excepción posible para la violación de la regla LIM superior en $e(d)$ (presión de vapor diaria, la cual está directamente relacionada con la humedad relativa), que se corresponde con la Ecuación (13) de rango rígido por la derecha ($0.8 < HR_{max}, HR_{min}, HR_{med} < 100$) para este trabajo; agregan que, sin embargo, para fines prácticos, el aire podría considerarse como saturado en determinados momentos del amanecer.

En esta prueba para la variable radiación global, en las diferentes estaciones de estudio los datos que no cumplieron con la condición de la Ecuación (18) son los siguientes: Agroeduca, 12 datos (0.55 %); El

Cardonal, 10 datos (0.46 %); El Fuerte, 2 datos (0.09 %); El Galomo, 16 datos (0.73 %); El Jardín, 16 datos (0.73 %); Santa Rosa I AC, 16 datos (1.08 %), y Chapingo, 14 datos (0.40 %).

Prueba de ceros. Para la prueba de ceros con la variable T_{min} , todas las estaciones del estudio valoraron como falsa la condición de la Ecuación (12), lo que significa que no hay datos marcados como sospechosos para esta prueba.

Schroeder *et al.* (2005) aplican esta prueba de ceros, así como las pruebas de rango, de paso, de persistencia, de instrumento y espacial para datos de alrededor de 40 estaciones del oeste de Texas, para periodos de observaciones a 5 y 15 minutos, y de 30 a 60 minutos; estos autores acumulan los errores de todas las pruebas y para la variable temperatura encontraron para las diferentes alturas desde el nivel del suelo: 1.5, 9 y 2 m, y errores de 0.003, 0.010 y 0.009 %, respectivamente.

De manera general, en la prueba de rango flexible, exceptuando Santa Rosa I para la variable radiación solar (23.25 %), la prueba de persistencia interna para la variable humedad relativa máxima en Santa Rosa I (77.43) y en El Galomo (Guanajuato) (44.36 %), las demás estaciones no presentaron datos espurios mayores al 20 % tal como reporta Araya-López (2011), en cuyo trabajo comenta que el porcentaje de datos atípicos de humedad relativa fue mayor al encontrado para datos de temperatura y nunca fue mayor al 20 % del total de datos de la red de estaciones que integran al Instituto Meteorológico Nacional de Costa

Rica; reporta asimismo porcentajes de datos atípicos para temperaturas máximas diarias de hasta 1.2 y 7.5 %, para temperaturas mínimas de 17.8 %; para humedad relativa mínima reporta datos atípicos de 4.3 y 8.7 %; y para humedad relativa máxima de 4.3 %.

En todas las pruebas aplicadas en este estudio se presentó algún dato marcado como sospechoso, espurio o inconsistente. La prueba de rango rígido para HR_{max} debe reconsiderarse al menos para Santa Rosa I AC, que es una zona de riego, ya que se puede alcanzar en días continuos HR_{max} de 100 %, como condición excepcional por lo mismo.

Finalmente, los datos analizados de acuerdo con la Figura 1, Figura 2, Figura 3 y Figura 4 presentan un comportamiento estacional y periódico, pero dado que se dispone de menos de 10 años de datos para todas y cada una de las estaciones objeto de estudio, no es posible observar una tendencia definida (se requiere de al menos 50 años de acuerdo con la OMM, 2007). De acuerdo con los reportes del IPCC (2001) es creciente para la temperatura y por lo tanto existiría un cambio en los límites extremos de la Ecuación (7) de este trabajo para cada estación, y entonces ello llevaría a una recalibración de estas ecuaciones.

Conclusiones

Los resultados de los procedimientos aplicados a datos históricos apoyan en determinar la confiabilidad de la información meteorológica con que cuenta una EMA y así establecer si se puede utilizar en el campo de interés, teniendo la confianza de que se va a llegar a inferencias correctas, o desechar esa estación o determinada variable. Además, indican qué sensores han trabajado dentro de su rango de precisión y también permite determinar datos sospechosos, no confiables o inconsistentes, lo cual ayuda a tomar la decisión de eliminarlos, reconsiderarlos o dejarlos marcados. Se recomienda implementar en tiempo real estos procedimientos, pues indicarán datos sospechosos que son más fáciles de verificar y a la vez definir si un sensor comienza a medir datos no confiables.

Referencias

- AENOR, Asociación Española de Normalización y Certificación. (2004). *Redes de estaciones meteorológicas automáticas: Directrices para la validación de registros meteorológicos procedentes de redes de estaciones automáticas (Validación en tiempo real)*. UNE 500540:2004. Madrid, España: Asociación Española de Normalización y Certificación.
- Allen, R. G., Smith, M., Pereira, L. S., & Perrier, A. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements. In: *FAO Irrigation and Drainage Paper No. 56*. Rome, Italy: Food and Agriculture Organization.
- Allen, R. G. (1996). Assessing integrity of weather data for reference evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, ASCE 122, 97-106. Recuperado de [https://doi.org/10.1061/\(ASCE\)0733-9437\(1996\)122:2\(97\)](https://doi.org/10.1061/(ASCE)0733-9437(1996)122:2(97))
- Allen, R. G. (2008). Quality assessment of weather data and micrometeorological flux - Impacts on evapotranspiration calculation. *Journal of Agricultural Meteorology*, 64(4), 191-204. DOI: 10.2480/agrmet.64.4.5
- Araya-López, J. L. (2011). Resultados de un control de calidad para datos de temperatura superficial de aire y humedad relativa. *Tecnología en Marcha*, 24, 33-49.



Araya-López, J. L., & Alfaro-Martínez, E. J. (2008). Prueba e implementación de algoritmos de control de calidad de datos de temperatura superficial del aire en un contexto operativo. *Tecnología en Marcha*, 21(3), 47-61.

ASCE-EWRI, American Society of Civil Engineers- Environmental & Water Resources Institute. (2005). *The ASCE Standardized Reference Evapotranspiration Equation. ASCE-EWRI Standardization of Reference Evapotranspiration Task Committee Report*. Recuperado de <http://www.kimberly.uidaho.edu/water/asceewri/>

Bulege-Gutiérrez, W., & Custodio, M. (2020). Cambio climático y retroceso glaciar en la cordillera Huaytapallana, Perú. *Tecnología y Ciencias del Agua*, 11(2), 229-261. DOI: 10.24850/j-tyca-2020-02-06

Campos-Aranda, D. F. (2020). Detección de registros homogéneos en 16 series amplias de precipitación anual del Altiplano Potosino, México. *Tecnología y Ciencias del Agua*, 11(3), 107-157. DOI:10.24850/j-tyca-2020-03-04

Estévez, G. J. (2008). *Diseño de un sistema integrado para el control de calidad de datos de estaciones meteorológicas automáticas* (disertación doctoral). Departamento de Agronomía, Universidad de Córdoba, España.

Estévez J., García-Marín, A. P., Morábito, J. A., & Cavagnaro, M. (2016). Quality assurance procedures for validating meteorological input variables of reference evapotranspiration in Mendoza province (Argentina). *Agricultural Water Management*, 172, 96-109. Recuperado de <https://doi.org/10.1016/j.agwat.2016.04.019>

Estévez, J., Gavilán, P., & Girálvez, J. V. (2011). Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*, 402, 144-154. DOI: 10.1016/j.jhydrol.2011.02.031

Feng, S., Hu, Q., & Qian, Q. (2004). Quality control of daily meteorological data in China, 1951-2000: A new dataset. *International Journal of Climatology*, 24, 853-870. DOI: 10.1002/joc.1047

Gleason, E. (2002). *Global daily climatology network, V 1.0*. Asheville, USA: National Climatic Data Center.

IPCC, Intergovernmental Panel on Climate Change. (2001). *Climate change: The scientific basis. Contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK, and New York, USA: Cambridge University Press.

Kubecka, P. (2001). A possible world record maximum natural ground surface temperature. *Weather*, 56, 218-221. Recuperado de <https://doi.org/10.1002/j.1477-8696.2001.tb06577.x>



- Majumdar, N. C., Mathur, B. L., & Kaushik, S. B. (1972). Prediction of direct solar radiation for low atmospheric turbidity. *Solar Energy*, 13, 383-394. DOI: 10.1016/0038-092X(72)90004-7
- Meek, D. W., & Hatfiel, J. L. (1994). Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology*, 69(1-2), 85-109. Recuperado de [https://doi.org/10.1016/0168-1923\(94\)90083-3](https://doi.org/10.1016/0168-1923(94)90083-3)
- Moradi, I. (2007). Quality control of global solar radiation using sunshine duration hours. *Energy*, 34(1), 1-6. DOI: 10.1016/j.energy.2008.09.006
- Obrien, K. J., & Keefer, T. N. (1985). Real-time data verification. Computer applications in water resources. In: *Procedures. ASCE Spec. Conf.* (pp. 764-770). Buffalo, New York, USA .
- OMM, Organización Meteorológica Mundial. (2007). *Función de las normales climatológicas en un clima cambiante. Programa mundial de datos y vigilancia del clima*. Ginebra, Suiza: Organización Meteorológica Mundial.
- Reek, T., Doty, S. R., & Owen, T. W. (1992). A deterministic approach to the validation of historical daily temperature and precipitation data from the Cooperative Network. *Bulletin of the American Meteorological Society*, 73, 753-762.

- Schroeder, J. L., Burgett, W. S., Haynie, K. B., & Sonmez, I. (2005). The West Texas Mesonet: A technical overview. *Journal of Atmospheric and Oceanic Technology*, 22, 211-222. Recuperado de <https://doi.org/10.1175/JTECH-1690.1>
- Shafer, M. A., Fiebrich, C. A., & Arndt, D. S. (2000). Quality assurance procedures in the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology*, 17, 474-494. Recuperado de <https://journals.ametsoc.org/doi/abs/10.1175/1520-0426%282000%29017<0474%3AQAPITO>2.0.CO;2>
- UNE 500540. (2004). *Redes de estaciones automáticas. Directrices para la validación de registros meteorológicos procedentes de redes de estaciones automáticas. Validación en tiempo real*. Madrid, España: Asociación Española de Normalización y Certificación.
- WMO, World Meteorological Organization. (1996). *Guide to Meteorological Instruments and Methods of Observation* (WMO-No. 8). Geneva, Switzerland: World Meteorological Organization.