

# Evaluación de la precisión de un algoritmo para la estimación de temperaturas medias basado en el método de gradiente altotérmico

## Evaluation of the accuracy of an algorithm for the estimation of average temperatures based on the high-temperature gradient method

Juan P. Demichelis<sup>1,2</sup>, Javier A. Carletto<sup>1,2</sup> y Jorge Leporati<sup>1</sup>

<sup>1</sup> Facultad de Ingeniería y Ciencias Agropecuarias (FICA) Universidad Nacional de San Luis

<sup>2</sup> Laboratorio de Energías Renovables (LER-FICA)

jpdemicheli@gmail.com, javiercarletto@gmail.com, jorgeleporati1560@gmail.com

### CientiFICA N°1

Año: 2024

pp. 23 – 34

### Historial del Artículo

**Recibido:**

15/04/2024

**Enviado revisión:**

17/04/2024

**Aceptado:**

30/04/2024

### Área temática:

Medio Ambiente

**Resumen:** Uno de los parámetros principales que se desean conocer en una determinada posición geográfica es la temperatura ambiente. Con el fin de estimar la temperatura en sitios donde no existen sensores, se diseñó y aplicó un algoritmo capaz de determinar la temperatura media mensual a partir de puntos de observación distantes conocidos. La Red de Estaciones Meteorológicas de la Provincia de San Luis, constituye un escenario ideal para la prueba del algoritmo. Para determinar su aplicabilidad se ensayó el algoritmo con 10.000 puntos aleatorios para determinar gráficamente la zona de aplicación. Para calcular el error se evaluaron 540 puntos donde el valor de temperatura era conocido en diferentes meses del año, determinando un error porcentual promedio del 6%. Se realizaron las evaluaciones estadísticas, las cuales arrojaron que las temperaturas reales vs estimadas son estadísticamente iguales con un 1% de significancia. Se pudo determinar que existen solo cuatro puntos ensayados con errores promedios mayores al 30%, coincidentes con estimaciones en zonas de extrema altura por lo que se continúa con el análisis de los resultados con el fin de identificar la confiabilidad para la aplicación del algoritmo en el territorio de San Luis, así como la factibilidad de mitigar el error.

### Palabras Claves:

Algoritmo, estimación de temperatura, valoración del error

### Keywords:

algorithm, temperature estimation, error assessment

**Abstract:** One of the main parameters of interest at a given geographical location is the ambient temperature. In order to estimate the temperature in places where there are no sensors, an algorithm capable of determining the monthly average temperature from known observation points was designed and applied. The REM - Network of Meteorological Stations of San Luis Province (Argentina), constitutes an ideal scenario for testing the algorithm. In order to determine its applicability, the algorithm was tested with 10,000 random points to graphically determine the application area. To calculate the error, 540 points were evaluated where the temperature value was known for different months of the year, determining a 6% average percentage error. Statistical evaluations were performed, which showed that the real versus estimated temperatures are statistically equal with 1% significance. It was determined that there are only four points tested with average errors greater than 30%, coinciding with temperature estimates in extreme altitude areas, so the analysis of the results continues to identify the reliability of the algorithm application throughout the San Luis territory, as well as the feasibility of mitigating the error.

## INTRODUCCIÓN

En la actualidad se cuenta con un número importante de puntos de observación con registros meteorológicos distribuidos sobre el territorio, sin embargo, y por distintas razones de aplicación, en algunas oportunidades es necesario conocer los parámetros climáticos en puntos geográficos alejados de los sensores instalados para lo cual se debe recurrir a métodos de estimación y ajuste a partir de datos conocidos en otros puntos.

Resulta indispensable el conocimiento del clima local en diversos aspectos tales como: definir estrategias de diseño de viviendas con el fin de lograr el confort humano sin depender de las energías contaminantes [1]; mejorar el confort en el interior de las viviendas dependiente de la interacción entre las condiciones interiores deseadas y las condiciones climáticas exteriores [2]; lograr un buen diseño de las construcciones con el uso pasivo de la energía solar dependiendo de la localización [3], etc.

Dentro de los parámetros climáticos, uno de los más importantes que es necesario conocer, y con cierto grado de exactitud, es la temperatura ambiente debido a la amplitud de sus aplicaciones, como por ejemplo cálculo de las cargas térmicas en el diseño de sistemas de refrigeración y aire acondicionado [4], estudios de enfermedades, pronósticos de epidemias, usos veterinarios, monitoreo medioambiental, estimación de las precipitaciones, pronósticos climáticos, evaluación de cambios climáticos, cálculo de evapotranspiración y modelación de rendimientos de los cultivos [5] entre otros.

Este parámetro comúnmente se mide a partir de observaciones sinópticas en estaciones meteorológicas que dependen de la distribución espacial e infraestructura instalada, la cual no siempre está diseñada para capturar un rango de datos confiables de una región de interés. La resolución espacial de estos datos puede ser compensada por medio de técnicas de interpolación entre sitios conocidos, pero cuando esta distribución no es suficiente por unidad de superficie, el resultado de la interpolación puede introducir errores inaceptables dependiendo de la aplicación [5],[6].

La estimación es importante también para conocer la variación espacial de la temperatura del aire, por ejemplo, en estudios relacionados con la agricultura a escala regional, como el balance hídrico y la estimación de evapotranspiración, ya que la baja densidad de estaciones meteorológicas no permite obtener dichos datos para una zona determinada [7].

Para realizar una estimación de temperatura con facilidad, y basada en un método para calcular las temperaturas mensuales promedio propuesto por De Fina y Sabella [8] se desarrolló un algoritmo capaz de entregar una estimación de temperatura en prácticamente cualquier punto que se encuentre rodeado por puntos de observación meteorológicos conocidos. El algoritmo fue implementado en MATLAB®, debido a su gran potencial para manejar grandes volúmenes de datos, y lograr una representación visual que permita fácil y rápidamente un análisis de la aplicación del método.

La red de estaciones meteorológicas de la provincia de San Luis, cuenta con 58 estaciones meteorológicas que disponen de datos en un mismo período de tiempo de 10 años, y que se encuentran distribuidas sobre el territorio, lo que genera un escenario propicio para ensayar el algoritmo. El objetivo del presente trabajo fue evaluar la exactitud y aplicabilidad del algoritmo para la estimación de la temperatura en sitios sin registro termométrico sobre el territorio de San Luis.

## METODOLOGÍA Y RESULTADOS

### Algoritmo de estimación

El algoritmo se basó en el método desarrollado por De Fina y Sabella [8] para la estimación de temperaturas en un punto determinado del territorio y su implementación fue descrita en detalle en [9]. El método consiste en calcular el gradiente térmico vertical mediano y con éste estimar las temperaturas en zonas sin observaciones termométricas. Este método permite determinar la temperatura promedio mensual en un punto con base en datos conocidos de cinco estaciones que lo rodean. El valor de temperatura se obtiene con precisión y confiabilidad lo que fue verificado por estudios realizados por Montoya y García [10].

### Funcionamiento del algoritmo

Para poder aplicar el algoritmo en un determinado punto, De Fina y Sabella [8] establecen el cumplimiento de cinco requisitos:

1. Observaciones de temperatura durante el mismo período. (Se tomaron temperaturas medias mensuales de la REM en un período de 10 años en todos los puntos de observación)
2. Diferencia de altitud apreciable. (Las estaciones de la REM se encuentran desde los 300 hasta los 1710 m.s.n.m.)
3. Una de las estaciones debe estar rodeada por las otras cuatro, de tal manera de constituir el centro de una equis, aproximadamente. (Logrado por la estrategia algorítmica Fig. 1)
4. La suma de las distancias de la estación central a las otras no debe ser mayor de 700Km.
5. La longitud de los brazos de la equis debe ser lo más uniforme posible.

Estos requisitos son satisfechos con los puntos de observación utilizados pertenecientes a la REM ya que el algoritmo selecciona la estación más próxima al punto incógnita donde traza una recta vertical y una horizontal separando el territorio en cuatro cuadrantes. Luego busca la estación más próxima en cada cuadrante con el fin de cumplimentar los requisitos y calcula el gradiente altotérmico, combinando de dos en dos los valores de temperatura y altitud de las 5 estaciones a utilizar, formándose un total de 10 combinaciones. Cada gradiente resulta de dividir la diferencia de temperatura entre dos estaciones, por su diferencia de altura en hectómetros. Estos gradientes se ordenan de mayor a menor o viceversa; considerando los signos, el promedio del 5º y 6º gradiente es el gradiente mediano de la zona considerada y será distinto para cada mes del año.

El algoritmo reduce las temperaturas (de la equis) al nivel del mar ( $t_{ANM}$ ), sumándole a la temperatura el producto del gradiente por su altura en hectómetros, tal como se observa en (1).

$$t_{ANM} = t_i + G \left( \frac{^{\circ}C}{hm} \right) \times H (hm) \quad (1)$$

Con estas temperaturas reducidas se unen las estaciones formando una equis como se observa en la Fig. 1. Cada lado se divide proporcionalmente a las diferencias de temperatura para trazar las isotermas rectas al nivel del mar. Fig. 2.

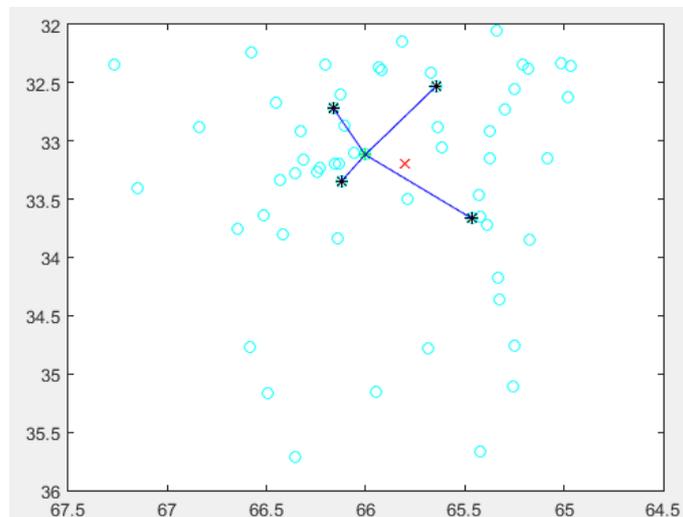


Figura 1. Equis formada por la unión de las estaciones conocidas alrededor del punto incógnita.

Una vez que se determina la temperatura estimada, esta será a nivel del mar, por lo que, para lograr una estimación de la temperatura en el punto de interés, solo interesa conocer la altura y realizar el cálculo según la Ec. (2).

$$t_i = t_0 - G \left( \frac{^{\circ}C}{hm} \right) \times H (hm) \quad (2)$$

Puede suceder que, por la geolocalización de las estaciones determinadas por el algoritmo, no sea posible trazar isotermas que alcancen el punto y por lo tanto no pueda estimarse la temperatura con los puntos de observación seleccionados. El algoritmo cambiará los puntos de observación conocidos utilizados e intentará

estimar la temperatura nuevamente. Aun cuando la densidad de puntos de observación sea alta, pueden existir lugares donde no pueda determinarse la estimación con este método, por lo que se procedió a realizar los ensayos descritos a continuación con el fin de determinar la aplicabilidad y el error de estimación.

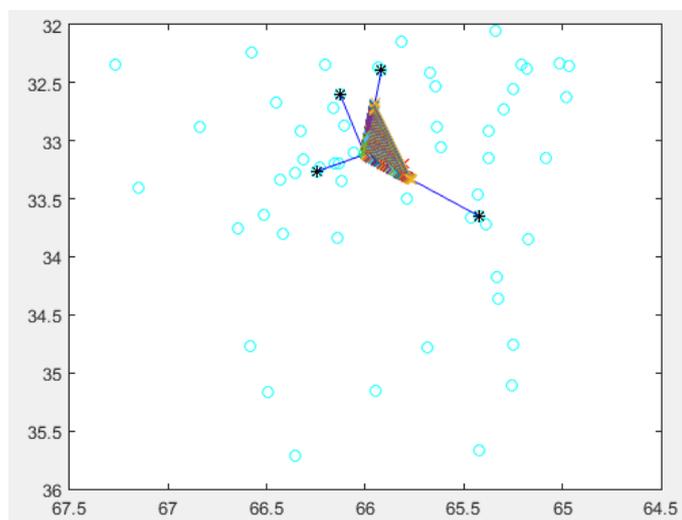


Figura 2. Isotermas a nivel del mar hasta alcanzar el punto incógnito

## Ensayo del algoritmo

### Datos utilizados para los ensayos

La provincia de San Luis cuenta desde el año 2007 con la REM - Red de Estaciones Meteorológicas, proyecto llevado adelante por el Gobierno provincial. Esta red brinda información del estado del tiempo en tiempo real y está compuesta por 74 puntos de observación entre estaciones meteorológicas automáticas y puntos de observación en diques y espejos de agua.

Tabla 1. Estaciones meteorológicas de la REM utilizadas. REF: referencia en la Figura 3

REF	Estación	REF	Estación	REF	Estación	REF	Estación	REF	Estación
1	Aerop. San Luis	13	El Amago	25	La Punilla	37	Navia	49	Tilisarao
2	Aerop. Valle del Conlara	14	El Durazno	26	La Punta	38	Nogolí	50	Unión
3	Alto Pelado	15	El Trapiche	27	La Toma	39	Nueva Galia	51	Valle de Pancanta
4	Anchorena	16	Estancia Grande	28	La Tranca	40	Paso Grande	52	Villa de Praga
5	Bajada Nueva	17	Fraga	29	Lafinur	41	Potrero de los Funes	53	Villa Gral. Roca
6	Baldecito	18	Justo Daract	30	Las Chacras	42	Queb. de las Higuieritas	54	Villa Larca
7	Batavia	19	La Angelina	31	Los Coros	43	San Francisco	55	Villa Mercedes
8	Beazley	20	La Botija	32	Luján	44	S.L. Rural	56	INTA VM
9	Buena Esperanza	21	La Calera	33	M. de Loyola	45	San Martín	57	Vla Reynolds
10	Concarán	22	La Cumbre	34	Merlo	46	San Miguel	58	Zanjitas
11	Coronel Alzogaray	23	La Esquina	35	Merlo Alto	47	Santa Rosa		
12	Desaguadero	24	La Florida	36	Naschel	48	Soven		

Para la prueba de nuestro algoritmo se tomaron en cuenta 58 estaciones pertenecientes a la REM que se encuentran ubicados estratégicamente en todo el territorio provincial (Tabla 1 y Fig. 3), cuyos datos estaban disponible y con un periodo de tiempo de acuerdo a lo solicitado por el método.

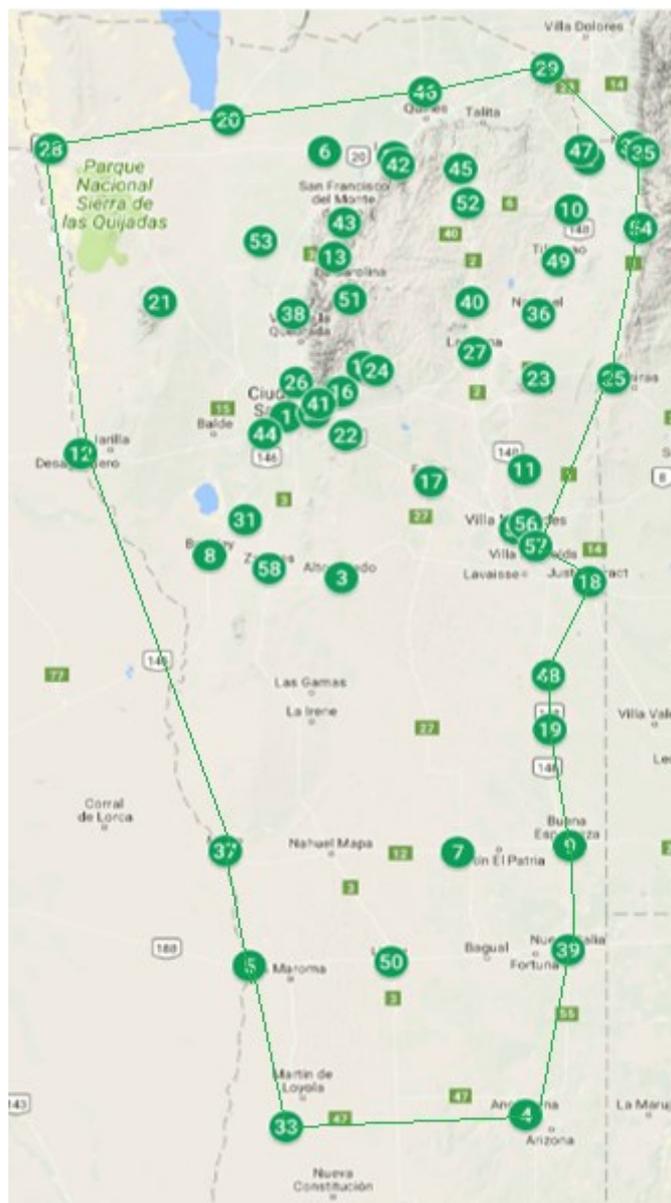


Figura 3. Distribución de estaciones meteorológicas utilizadas sobre la provincia de San Luis. Polígono: representa el límite máximo de funcionamiento del algoritmo.

Los datos provenientes de estos puntos de observación fueron procesados logrando una base de datos consistente para la prueba del algoritmo.

### Aplicabilidad del algoritmo

Debido al funcionamiento del método, el algoritmo será capaz de estimar una temperatura en puntos que se encuentren obligatoriamente dentro del polígono determinado por los puntos de observación exteriores a la zona estudiada. Dicho de otro modo, el algoritmo funciona dentro de los límites impuestos por las propias estaciones conocidas (Fig. 3). Este polígono irregular dentro del cual funciona el algoritmo cambia, si incorporamos o quitamos puntos de observación.

Aun estando dentro del polígono, es necesario entender las zonas de aplicabilidad del algoritmo. Para ello se ensayó el mismo con 10.000 puntos con coordenadas geográficas aleatorias sobre un rectángulo que contiene a la provincia de San Luis.

La Fig. 4, muestra los puntos ensayados, identificándose con cruces azules a los puntos donde el algoritmo estimó un valor, y con equis rojas, los puntos donde el algoritmo no pudo realizar la estimación.

Superponiendo los puntos en un mapa de San Luis, puede observarse que el algoritmo pudo determinar un valor estimado de temperatura en la mayoría de los puntos interiores al polígono. Esto también se debe al número y distribución de las estaciones meteorológicas que se dispone en el territorio.

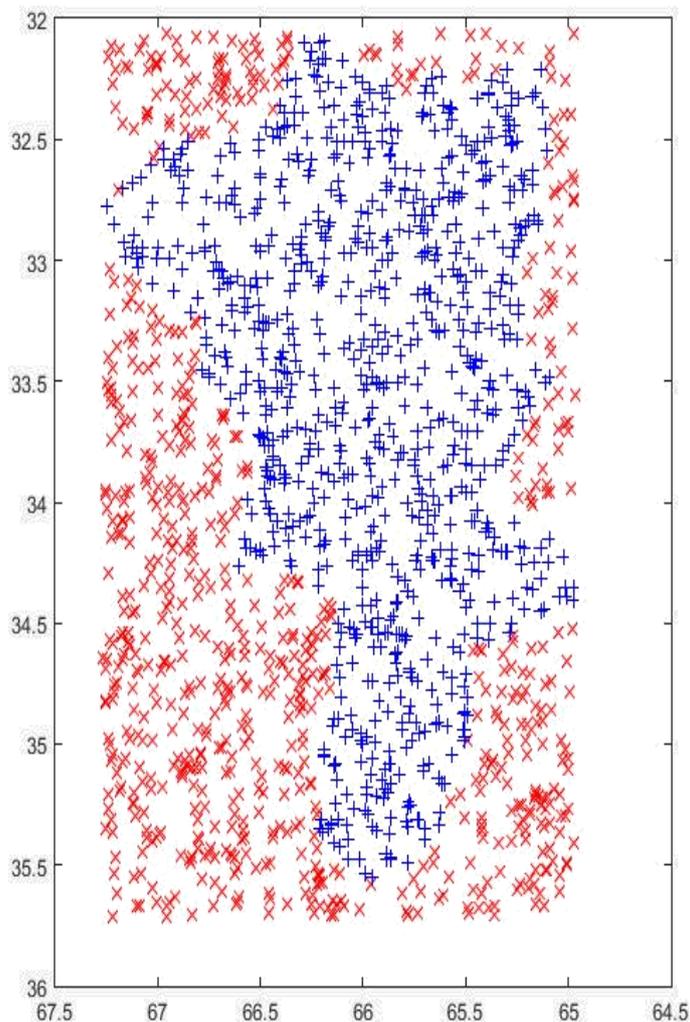


Figura 4. Puntos en los que se ensayó la aplicabilidad del algoritmo. Azules: Se pudo estimar la temperatura. Rojos: no se pudo estimar.

## Determinación del error de estimación

Si bien la metodología descripta permite determinar aproximadamente la zona de aplicabilidad según los datos conocidos, no permite determinar con qué nivel de precisión se realiza la estimación.

Para poder determinar el error de estimación es necesario contar con puntos de prueba con un historial de datos, por lo que se utilizó la siguiente estrategia:

Se calcula para todos los meses del año, el valor de temperatura promedio en las coordenadas de los puntos de observación de la Red de Estaciones Meteorológicas de la provincia de San Luis, retirando esa estación de la base de datos y considerándola como valor verdadero de la temperatura en ese punto.

La cantidad de posibilidades a evaluar se puede calcular como la cantidad de estaciones (58) multiplicado por la cantidad de meses (12). De estas 696 posibilidades totales, se descartaron las correspondientes a las estaciones 4, 5, 9, 12, 20, 25, 28, 29, 33, 34, 37, 39, 54 situadas en los vértices del polígono representado en la Fig. 3 donde por definición del método, no se puede realizar la estimación de temperatura, resultando 540 posibilidades de ensayo del algoritmo.

El algoritmo arrojó un resultado en 459 puntos, es decir que pudo determinar un valor estimado de temperatura en el 85% de las coordenadas de los puntos conocidos.

Para cada una de estas posibilidades se calculó el error en la estimación en función de la Ec. (3), arrojando un error relativo porcentual promedio del 6%, cuyo análisis estadístico comparativo se observa en la siguiente sección.

$$ERP = \frac{\text{Temperatura Medida} - \text{Temperatura Estimada}}{\text{Temperatura Medida}} * 100 \quad (3)$$

La Tabla 2, muestra el error relativo porcentual en valor absoluto, con una escala de colores que cuantifican el error.

Para comprender estadísticamente la precisión de la estimación arrojada por el algoritmo, se realizó un análisis estadístico descriptivo y una comparación entre las temperaturas reales y estimadas utilizando el software R en su versión 4.0.2.

En 4 puntos ensayados los errores promedios de estimación superaron el 30%, lo que podrían resultar inadmisibles para ciertos usos.

Estas posiciones corresponden a las estaciones ubicadas en altitudes extremas, situándose en la zona serrana, donde los puntos de observación conocidos que utilizó el algoritmo probablemente no sean los adecuados.

Los cuatro puntos que arrojan errores inadmisibles se corresponden con los puntos de observación localizados en San Martín (45), a 930 m.s.n.m. donde la estimación produjo un error relativo porcentual promedio de 31%; El Amago (13) con una altura de 1710 m.s.n.m. y un error promedio de 102%; Valle de Pancanta (51) con una altura 1598 m.s.n.m. y un error del 48,2% y Villa Praga (52) a 994 m.s.n.m. con un error del 27% (Fig. 3). Estos cuatro puntos no se tuvieron en cuenta para el análisis estadístico ya que fueron considerados como valores atípicos.

## Análisis estadístico

En Fig. 5, el gráfico de diagrama de caja muestra que no se observa una diferencia significativa entre la temperatura real y la temperatura estimada. Se puede apreciar una leve discrepancia en las barras, pero no se identifican valores atípicos en los datos.

Para validar lo expresado anteriormente, se aplicó la prueba paramétrica para dos muestras apareadas o relacionadas. Esta técnica, se basa en comparar, si las medias de dos grupos, son iguales o semejantes, o si la diferencia entre las medias de las dos temperaturas es igual a cero; para ello necesitamos que se cumplan el supuesto de normalidad de los datos.

Tabla 2. Valor absoluto del error relativo porcentual para las posiciones conocidas. NC: No Calcula

Estación	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Agos.	Sept.	Oct.	Nov.	Dic.
1	3,4%	19,7%	21,1%	16,3%	15,0%	28,6%	9,0%	3,7%	11,7%	18,4%	12,0%	0,9%
2	1,3%	3,6%	0,8%	6,1%	4,1%	0,3%	3,1%	7,7%	2,6%	9,0%	4,6%	0,7%
3	11,4%	11,2%	10,6%	4,1%	13,6%	18,2%	14,2%	18,8%	15,7%	11,5%	3,8%	10,8%
6	0,5%	0,1%	2,2%	6,1%	9,9%	18,4%	16,5%	10,7%	0,9%	5,7%	0,8%	1,6%
7	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
8	NC	NC	NC	8,6%	5,5%	24,2%	NC	NC	7,9%	1,0%	10,4%	NC
10	2,0%	1,8%	2,4%	4,8%	2,4%	1,9%	3,2%	3,1%	3,1%	4,4%	2,8%	2,4%
11	2,4%	0,6%	5,1%	3,7%	0,2%	14,9%	4,7%	3,6%	7,7%	7,5%	6,8%	0,2%
13	95,0%	100,9%	76,4%	118,6%	115,4%	137,1%	70,0%	92,3%	100,5%	129,5%	116,6%	77,6%
14	4,3%	4,3%	0,9%	1,1%	4,7%	8,3%	6,7%	7,0%	2,9%	1,1%	2,3%	5,3%
15	1,9%	1,2%	2,9%	6,8%	10,2%	29,5%	22,8%	14,1%	9,1%	3,5%	4,8%	2,8%
16	0,9%	0,8%	4,4%	5,1%	1,7%	0,2%	4,1%	0,8%	4,1%	5,6%	3,5%	0,3%
17	5,9%	5,0%	3,7%	5,8%	3,3%	24,4%	6,6%	3,3%	6,7%	8,1%	7,1%	5,3%
18	NC	NC	NC	0,1%	3,3%	5,2%	0,5%	NC	NC	1,5%	NC	NC
19	0,5%	0,5%	1,8%	0,1%	0,9%	2,5%	2,7%	3,4%	3,1%	14,7%	10,3%	1,4%
21	13,6%	9,0%	9,4%	9,7%	13,7%	NC	15,3%	0,7%	12,5%	7,2%	9,6%	10,9%
22	13,2%	19,4%	10,8%	5,8%	10,8%	8,3%	6,8%	29,0%	13,3%	10,3%	13,0%	11,1%
23	12,6%	16,6%	11,2%	18,4%	16,9%	17,3%	10,2%	10,9%	14,1%	15,6%	15,1%	16,9%
24	5,7%	6,7%	9,2%	14,0%	17,2%	27,0%	27,5%	17,2%	13,0%	11,1%	9,9%	6,2%
25	NC	NC	NC	NC	NC	4,8%	NC	10,6%	11,4%	NC	NC	NC
26	1,7%	0,6%	5,4%	9,7%	3,6%	10,7%	8,6%	2,1%	6,1%	10,5%	8,1%	2,2%
27	10,3%	11,7%	7,8%	11,1%	11,1%	3,9%	3,5%	3,6%	10,6%	16,0%	15,0%	13,6%
30	1,5%	0,6%	1,1%	2,2%	5,2%	28,7%	34,2%	20,1%	32,5%	8,9%	5,2%	1,5%
31	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
32	2,2%	5,6%	4,9%	11,8%	5,7%	4,0%	4,7%	2,1%	0,8%	7,4%	6,7%	1,3%
35	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
36	0,5%	0,4%	0,8%	0,2%	11,9%	31,2%	13,6%	14,8%	1,3%	0,3%	0,3%	1,1%
38	8,5%	8,7%	8,9%	11,2%	14,9%	20,1%	22,5%	18,0%	15,9%	13,1%	11,4%	9,0%
40	2,4%	2,1%	2,3%	2,8%	2,2%	1,6%	1,7%	0,2%	3,5%	0,6%	1,7%	4,2%
41	2,0%	1,1%	4,8%	1,8%	5,9%	11,7%	10,9%	2,8%	4,3%	5,0%	5,7%	1,3%
42	0,8%	3,6%	0,9%	8,3%	0,5%	6,9%	7,1%	4,9%	7,8%	2,9%	3,0%	2,4%
43	22,2%	24,2%	24,0%	26,8%	30,2%	32,4%	32,0%	8,4%	31,1%	27,9%	3,6%	22,1%
44	17,4%	38,5%	50,9%	49,8%	31,3%	27,1%	38,5%	27,8%	30,5%	27,6%	20,2%	0,2%
45	17,2%	14,9%	20,6%	20,7%	40,7%	22,2%	75,7%	51,8%	45,8%	24,1%	19,1%	20,2%
46	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
47	NC	NC	NC	6,9%	NC	NC	NC	NC	NC	7,8%	NC	NC
48	1,6%	0,1%	1,7%	4,2%	3,3%	5,2%	0,1%	0,6%	4,3%	25,8%	13,8%	2,2%
49	2,0%	2,7%	2,7%	2,4%	2,1%	2,7%	2,2%	1,6%	2,0%	2,4%	2,1%	2,8%
50	3,7%	0,0%	5,6%	5,4%	9,1%	2,8%	0,8%	2,1%	2,9%	2,6%	0,8%	0,1%
51	34,0%	33,3%	34,1%	41,4%	52,6%	68,1%	51,1%	70,1%	67,5%	52,0%	38,6%	35,2%
52	15,6%	16,0%	17,8%	19,5%	29,8%	50,8%	44,0%	37,5%	34,7%	22,0%	17,2%	18,9%
53	1,4%	0,7%	0,1%	0,1%	0,6%	10,7%	13,7%	11,5%	11,3%	8,6%	6,6%	0,9%
55	8,0%	6,6%	5,9%	3,0%	6,1%	13,5%	15,2%	10,0%	13,1%	7,4%	4,2%	6,2%
56	12,2%	7,0%	10,1%	11,9%	14,6%	23,1%	18,4%	13,3%	16,7%	11,5%	11,5%	8,4%
57	9,2%	1,9%	9,1%	17,2%	18,2%	36,1%	16,5%	8,7%	15,2%	8,9%	12,4%	5,8%
58	4,7%	6,6%	7,8%	7,6%	5,5%	4,6%	5,9%	10,6%	0,1%	7,7%	5,4%	4,0%

### Box Plot para Temperatura Real y Temperatura Estimada

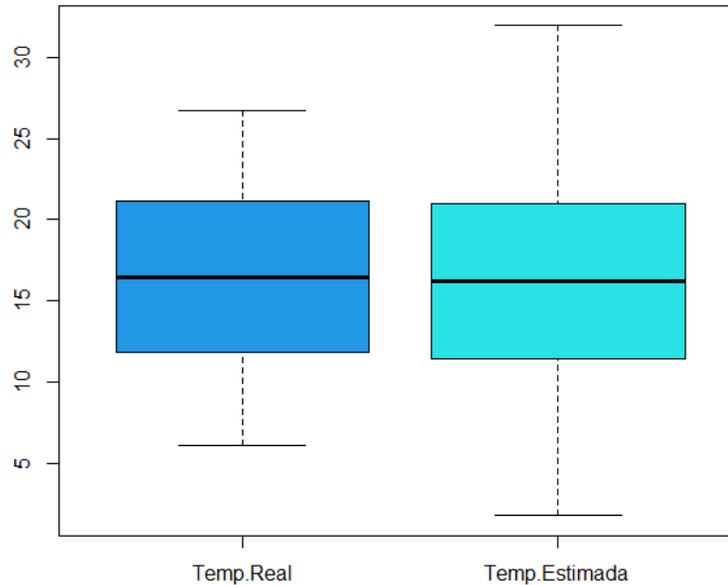


Figura 5. Grafica comparativa de los puntos ensayados y la temperatura real.

### Normalidad

Para determinar la normalidad de las muestras se aplicó el test de Shapiro-Wilks. Considerando un nivel de significancia del 1%, el test arrojó un p.valor =  $2.919e-09 < 0.01$ , para la temperatura real y un p.valor =  $0.0002242 < 0.01$  para la temperatura estimada, indicando de esta manera que los datos no son normales, lo que puede verse gráficamente, en las Fig. 6 y Fig. 7 respectivamente.

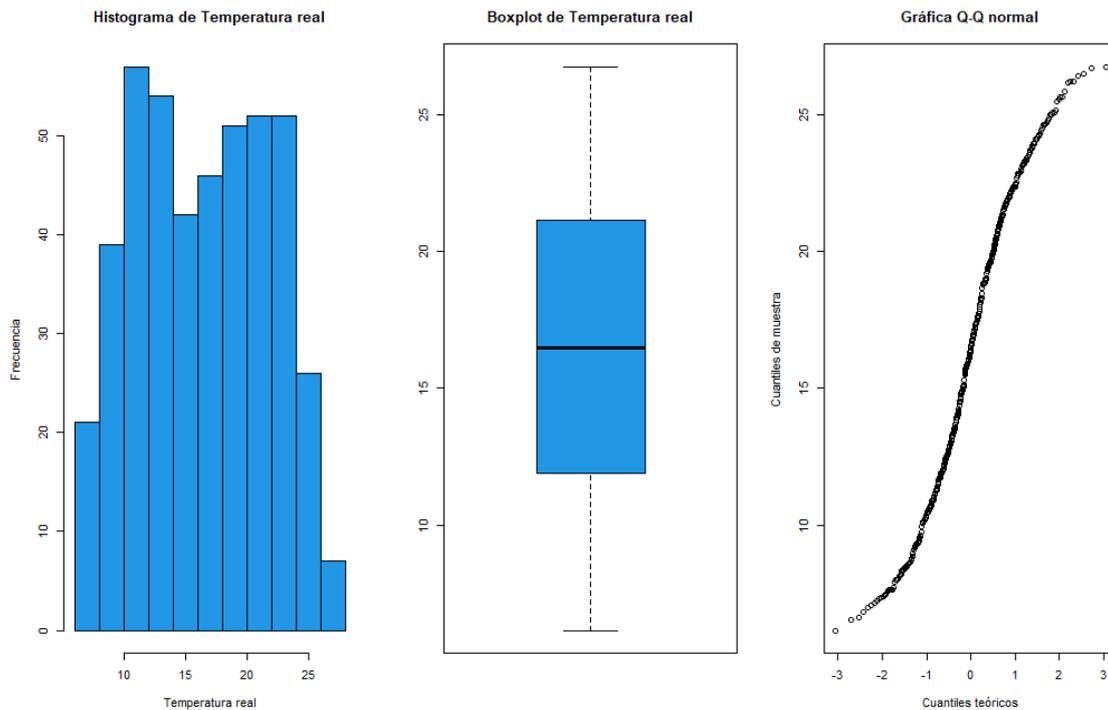


Figura 6. Graficas de normalidad para la temperatura real.

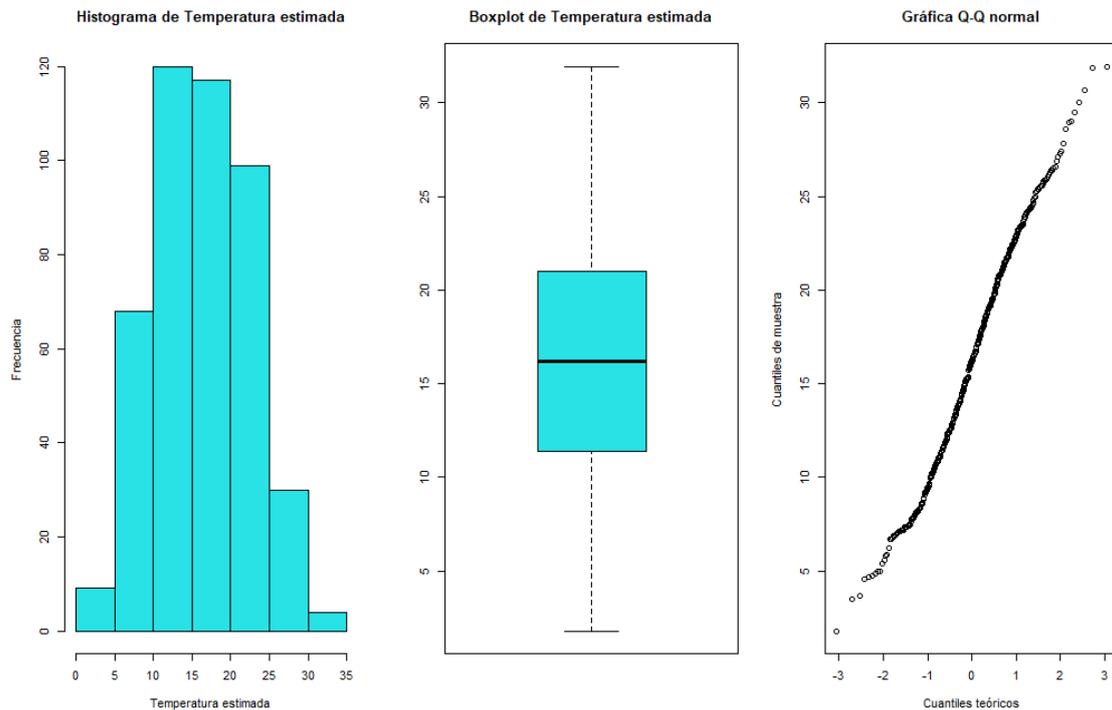


Figura 7. Graficas de normalidad para la temperatura estimada.

### Alternativa robusta

Debido al no cumplimiento de los supuestos de normalidad estudiado anteriormente, se opta en aplicar una técnica robusta, como alternativa de la prueba T para muestras apareadas, que es el test de Yuen.

Este test se fundamenta en que, dado que los datos rara vez son normales, muchos métodos estadísticos clásicos asumen distribuciones normalmente distribuidas, especialmente cuando se trata de muestras pequeñas. Paralelamente, para muestras grandes, el teorema del límite central expone que no es un punto importante a tener en cuenta. Cuando se trata de pruebas estadísticas donde existen desviaciones de normalidad prominentes de los datos tales como datos muy sesgados, datos con valores atípicos y distribución de colas pesadas, es necesario la aplicación de las técnicas específicas para estos casos, ya que las violaciones de la normalidad influyen en los resultados de las pruebas estadísticas clásicas. Los métodos robustos para la estimación y las pruebas estadísticas robustas brindan, ante la falta de normalidad, una excelente opción para el análisis. Históricamente, los primeros desarrollos se remontan a los años 60 con publicaciones de Tukey [11], Huber [12] y Hampel [13]. El objetivo principal de estos métodos, es producir resultados que sean menos sensibles a pequeñas desviaciones del supuesto funcional de forma, Yuen es uno de ellos y se utiliza para comparar muestras independientes o apareadas basándose en medias recortadas [14].

### Aplicación del test de Yuen

El resultado obtenido con el test fue satisfactorio. Considerando un nivel de significación del 1%, el test arrojó un  $p\text{-valor} = 0.01263 > 0.01$  (Fig. 8), por lo que no se rechaza la hipótesis de medias semejantes entre las dos muestras, indicando que la media de la temperatura real es la misma o similar que la media de la temperatura estimada.

```
> yuend(Temp.Real,Temp.Estimada,paired=T)#test Robusto
Call:
yuend(x = Temp.Real, y = Temp.Estimada, paired = T)

Test statistic: 2.5109 (df = 268), p-value = 0.01263

Trimmed mean difference: 0.27061
95 percent confidence interval:
0.0584      0.4828

Explanatory measure of effect size: 0.03
```

Figura 8. Resultado de la alternativa robusta de Yuen.

### Resumen de parámetros estadísticos

Si analizamos los valores de medias, medianas y desvíos estándar; podemos apreciar que la diferencias entre los valores de temperatura real y temperatura estimada son prácticamente despreciables. (Tabla 3)

Tabla 3. Parámetros estadísticos de la Temperatura Real y la Estimada.

Parámetro Estadístico	Temperatura Real	Temperatura Estimada
Medias Estimadas	16.17650	15.57256
Medianas Estimadas	16.17000	15.79281
Desvíos Estimados	5.415624	6.655433

## CONCLUSIONES

El estudio realizado permitió evaluar de forma satisfactoria la exactitud y aplicabilidad del algoritmo para la estimación de la temperatura en sitios sin registro termométrico sobre el territorio de San Luis. El algoritmo implementado basado en el método del gradiente altotérmico funciona correctamente con una aplicabilidad alta, del orden del 85%, y errores aceptables promedios del 6%.

Debido al número y distribución de las estaciones meteorológicas que se dispone en el territorio, el algoritmo fue capaz de determinar la temperatura en prácticamente cualquier parte del territorio enmarcado por un polígono irregular trazado mediante la localización de los puntos de observación más alejados, siendo menor la probabilidad de cálculo en las zonas próximas a las aristas de dicho polígono.

Para las muestras analizadas, las Temperaturas Reales vs. las Temperaturas Estimadas, son estadísticamente similares al 1% de significancia, lo que se pudo demostrar gráfica y analíticamente.

No obstante, se detectaron e identificaron puntos localizados en altura extrema donde los errores mayores al 30% en promedio, podrían resultar inaceptables para ciertos usos por lo que se requerirá un ajuste del algoritmo. Probablemente los puntos conocidos utilizado por el algoritmo no cumplirían las características ideales para la aplicación del método, o el método necesite ser ajustado para estas situaciones especiales, por lo que será necesario continuar el análisis de aplicabilidad para mitigar estos errores.

## REFERENCIAS

- [1] Gatani M., Bracco M., Angiolini S., Jerez L., Pacharoni A., Sánchez G., Tambussi R., Avalos P. (2008) Definición de indicadores de análisis de diseño sustentable. El caso de una vivienda serrana en Córdoba. *Actas de la XXXI Reunión de Trabajo de la Asociación Argentina de Energías Renovables y Ambiente - ASADES*
- [2] Romero, F.; Galimberti, P.; Ziletti, M.; Paisio, G.; Soler, L. (2009). Progresos en la sistematización de variables climáticas para la ciudad de Río Cuarto. *Avances en Energías Renovables y Medio Ambiente. Vol. 13. 11.77-11.83.*

- [3] Thomas L. P. y Marino B. M. (2008). Comportamiento térmico de una vivienda frente a variaciones importantes de temperatura en verano. *Avances en Energías Renovables y Medio Ambiente*, Vol. 12, 2008. ISSN 0329-5184
- [4] Rosario Moreno, L.A.; Pereyra E. J.; Vielma J. C. (2001). Estimación de temperatura y humedad relativa en Venezuela mediante redes neuronales. *Ciencia e Ingeniería Vol 22, Núm. 1*
- [5] Méndez Jock, A., Venus, V. (2006). Estimación de la temperatura ambiental a escala regional mediante procesamiento de imágenes de satélites. *Ciencias de la Tierra y el Espacio*, Vol.7, pp.10--20, ISSN 1729-3790
- [6] Lennon, J. L., Turner, J.R.G. (1995). Predicting the spatial distribution of climate: Temperature in Great Britain. *Journal of Animal Ecology [J. ANIM. ECOL.]*, vol. 64, no. 3, pp. 370-392, 1995. Recuperado de <http://www.csa.com/htbin/linkabst.cgi?issn=0021-8790&vol=64&firstpage=370>
- [7] Flores F. y Lillo M. (2010). Simple air temperature estimation method from modis satellite images on a regional scale Chilean. *Journal of agricultural research* 70(3):436-445
- [8] De Fina, A.; Sabella, I. J. (1959). Cálculo de temperaturas medias en localidades montañosas carentes de observaciones termométricas. *Revista de la Facultad de Agronomía, La Plata, Argentina* 35(2): 127-145
- [9] Carletto, J. y Demichelis, J. P. (2018) Algoritmo en versión beta para la estimación de temperaturas promedio mensuales de un punto de la Provincia de San Luis basado en puntos de observación conocidos. *Acta de la XLI Reunión de Trabajo de la Asociación Argentina de Energías Renovables y Medio Ambiente Vol. 6, pp. 06.63-06.74. ISBN 978-987-29873-1-2*
- [10] Montoya, M., J. M., y Garcia B., J. (1971). Comparación de dos técnicas para estimar temperaturas medias en localidades carentes de registros con fines agroecológicos. *Turrialba* 21(1):112-115.
- [11] Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In I. Olkin (Ed.), *Contributions to Probability and Statistics*, (pp. 448-485). Palo Alto: Stanford University Press.
- [12] Huber, P.J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35, 73-101.
- [13] Hampel, F. R. (1968). *Contributions to the Theory of Robust Estimation*. PhD Thesis, University of California, Berkeley.
- [14] Mair, P., Wilcox, R. (2020) Robust statistical methods in R using the WRS2 package. *Behav Res* 52, 464-488. <https://doi.org/10.3758/s13428-019-01246-w>