DOI: 10.24996/ijs.2023.64.3.44

ISSN: 0067-2904

Treatment Missing Data of Air Temperature in Iraq by Using Mean Method

Ali Hamid Yassen[1], Asraa Khtan Abdul Kareem [2]
[1] Weather Forecasting Department, Iraqi Meteorological Organization and Seismology, Baghdad, IRAQ.
[2] Department of Atmospheric Sciences, College of Science, Mustansiriyah University, Baghdad, IRAQ.

Received: 24/4/2022 Accepted: 15/8/2022 Published: 30/3/2023

Abstract

Frequent data in weather records is essential for forecasting, numerical model development, and research, but data recording interruptions may occur for various reasons. So, this study aims to find a way to treat these missing data and know their accuracy by comparing them with the original data values. The mean method was used to treat daily and monthly missing temperature data. The results show that treating the monthly temperature data for the stations (Baghdad, Hilla, Basra, Nasiriya, and Samawa) in Iraq for all periods (1980-2020), the percentage for matching between the original and the treating values did not exceed (80%). So, the period was divided into four periods. It was noted that most of the congruence values increased, reached in summer (70%-100%), and decreased somewhat in winter. While the daily treatment using the mean method for the stations Baghdad and Basra (2010-2020), it turns out that most of the congruence values in the summer ranged (70%-100%), but in winter, the congruence is often decreased. Therefore, this method gives high accuracy when treating monthly and daily temperatures in summer and less in winter.

Keywords: missing data, treatment, temperature, mean method, Iraq.

معالجة البيانات المفقودة لدرجة الحرارة في العراق باستخدام طريقة المتوسط

علي حامد ياسين[1]*, اسراء قحطان عبدالكريم[2]

[1]قسم التنبؤ الجوي, الهيئة العامة للانواء الجوية والرصد الزلزالي, بغداد, العراق

[2]قسم علوم الجو, كلية العلوم, الجامعة المستنصرية, بغداد, العراق

الخلاصة

تعد البيانات المستمرة في سجلات الطقس ضرورية للتنبؤ وتطوير النماذج العددية والبحوث ، ولكن قد تحدث انقطاعات في تسجيل البيانات لأسباب مختلفة. لذلك ، تهدف هذه الدراسة إلى إيجاد طريقة لمعالجة هذه البيانات المفقودة ومعرفة دقتها من خلال مقارنتها بقيم البيانات الأصلية. تم استخدام طريقة المتوسط لمعالجة بيانات درجة الحرارة المفقودة اليومية والشهرية. بينت النتائج ان معالجة بيانات درجات الحرارة الشهرية للمحطات (بغداد والحلة والبصرة والناصرية والسماوة) في العراق لجميع الفترة (1980−2020) لم تتجاوز نسبة المطابقة بين القيم الاصلية والمعالجة (80٪). لذلك ، تم تقسيم الفترة إلى أربع فترات. لوحظ أن معظم قيم

* Email: alihamid.1987@yahoo.com

التطابق قد زادت حيث وصلت في الصيف (70٪ –100٪) وتنخفض نوعاً ما في الشتاء. اما في المعالجة اليومية باستخدام طريقة المتوسط لمحطتي بغداد والبصرة (2010–2020) اتضح ان معظم قيم التطابق في الصيف تراوحت بين (70٪ –100٪) ولكن في الشتاء يكون التطابق غالبا اقل. لذلك تعطي هذه الطريقة دقة نتائج عالية عند معالجة درجات الحرارة الشهرية واليومية في الصيف وأقل في الشتاء.

1. Introduction

The temperature directly impacts our lives, which is one of the most significant components of meteorological and climatological materials [1]. The temperature ranges determine whether the weather is cold or hot [2]. The air's temperature is an essential component that influences the weather and directly relates to other meteorological characteristics, such as solar radiation, air humidity, and atmospheric pressure [3]. The world meteorological organization recommended taking temperature readings from 1.25 to 2 m above ground level using a weather station screen thermometer [4], where the information was gathered in synoptic data. Climate data is the formal data register, usually supplied after conducting quality control [5]. Climate studies require continuous time series data for climate studies and analysis. It is possible to lose some data, but approaches to filling gaps in the data series are required in climate change prediction models and research investigations [6]. Through time, weather stations may begin to be out of service or stopped for repair; through those periods, there will be missing data [7] or for a reason. Problems in storing or transferring data due to sensor failure [8] and surface weather station networks are possible in many parts of the world because of geography's composition, maintenance costs, and complexity [8]. Many statistical techniques are available to find missing temperature data, including deletion, direct estimation, and imputation [9]. Nassir et al. (2018) used the Autoregressive integrated moving average (ARIMA) model for estimating the missing data of air temperature, relative humidity, and wind speed for mean monthly variables in different time series at three stations in Iraq, Sinjar, Baghdad, and Al-Hai. The results of the ARIMA model are accurate, but data should be available in sufficiently large numbers to estimate the missing data within the time series [10]. Jaber Rahimi et al. (2017) used the Cumulative Distribution Function (CDF) method to complete a daily minimum missing temperature data series of highly elevated stations in Iran from 1965 to 2010 [11]. Marina Popolizio et al. (2021) used Generative Adversarial Imputation Networks (GAIN) to complete temperature time series, and the obtained results were very encouraging [12]. Most of the studies in Iraq used the mean method for the ease and speed of presenting results, so this study aims to appropriately use treatment for the missing temperature data.

2. Study area and Dataset

Iraq's weather is primarily dry and semi-arid with high temperatures, often on the same day, day and night, and summer and winter, and this great contradiction is clear from the climatic extremism that symbolizes the state of the continental climate in general [13]. In summer, the temperature in the middle and southern sections of the country exceeds 50 C, while in winter, the temperature in the north and mountain areas can drop below freezing [13]. The monthly mean temperature varies from 30-45 C^o in summer to 5-20 C^o in winter. In the south of Iraq, the climate is characterized as dry, whereas, in the north, it is classified as semi-arid [14]. The mean monthly temperature was obtained from the Iraqi Meteorological and Seismological Organization (IMOS) for 40 years (1980-2020) and the mean daily temperature for 11 years (2010 -2020). Data were collected from five weather stations in the center and south of Iraq. The central stations were Baghdad and Hilla, located between latitudes (32° 27' and 33° 18` N) north of the equator and longitudes (44° 24' and 44° 27` E) east of Greenwich. Also, some stations in southern Iraq (Basra, Nasiriya, Samawa), which are geographically located in the northern hemisphere between latitudes (30° 31' and 31° 16` N) north of the equator and longitudes (45° 16'-47°47` E) east of Greenwich, as shown in Table (1) and Figure (1).

Table 1: Locations of weather stations used in the study [15]

Station	Station Number	Longitude	Latitude	Elevation
Baghdad	650	44° 24`	33° 18`	32 m
Hilla	657	44° 27`	32° 27`	27 m
Basra	689	47° 47`	30° 31`	2 m
Nasiriya	676	46° 14`	31° 01`	5 m
Samawa	674	45° 16`	31° 16`	11 m

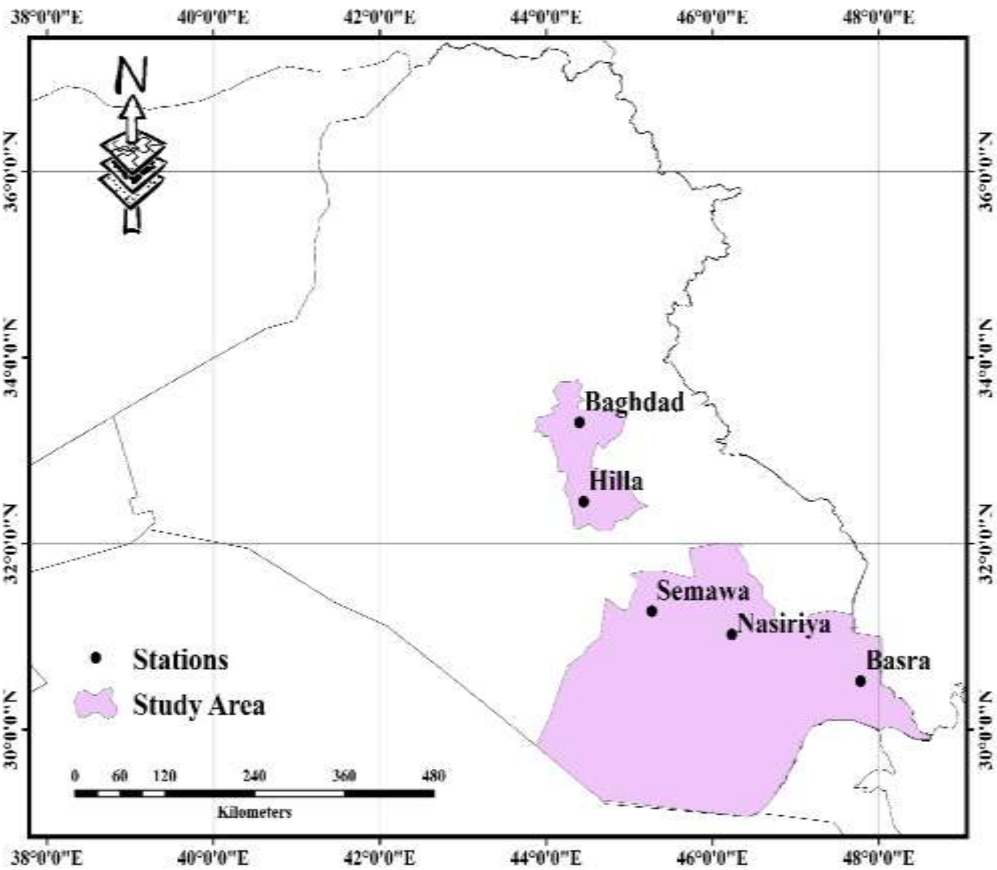

Figure 1: Locations of weather stations used in the study

3. Methodology
3.1 Mean Method

The mean is a value around which the values of a group are grouped and through which the rest of the group values can be evaluated, so the values are the arithmetic mean. It was calculated by summing and dividing the total by the number of values [16].

$$\bar{x} = \frac{\sum x}{n} \qquad \text{--------1}$$

Where $\bar{x}$ = mean of values, $\sum x$ = the sum of the average daily or monthly recorded temperature, n = number of days or months.

The arithmetic mean method was used to find the congruence in temperature between the actual values and the mean for the study area, where the study included two parts, daily and monthly.

The first daily treatment: taking the mean daily temperature every month and every year for the daily period for Baghdad and Basra stations and then subtracting it from the daily temperature values for the same month. The values are considered identical if the results range between (0-2.5) °C depending on the standard deviation values shown in Table (2-A, B). However, if the values are more significant than (2.5) °C, they are considered values that are not close to the actual values, so we exclude and extract the difference from them (the percentage of difference).

The second monthly treatment: the general mean of each month was taken for the period (1980 - 2020). A variance was observed in the mean for the entire period, so the study period was divided into four periods (every ten years) for the stations (Baghdad, Hilla, Basra, Nasiriya, and Samawa), where the comparison was made between the value of the general mean for each period and the values of the original period. The values are considered identical if the difference ranges between 0-1.5 °C (based on standard deviation values), while if it is higher than 1.5 °C, the values are isolated, and the percentage difference is calculated.

Table (2- A): The daily mean temperature for the daily period and the standard deviation of the Baghdad stations.

Baghdad	January		February		March		April		May		June		July		August		September		October		November		December	
	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
2010	13.8	2.8	15	4.2	19.5	4.3	24	2.3	30.5	3.4	33.1	2.2	37.3	1.7	37.7	1.6	33	1.6	27	2.9	18	2.2	13	2
2011	10.2	2	12.7	4.5	17.5	3.2	23.6	2.9	29.5	3.5	34.3	2	37	2	35.7	1.6	31.4	1.6	23.6	3.4	13.7	3.8	9.7	2.3
2012	9.5	2.6	12	2.5	15.4	3.7	25.2	2.4	30.6	2.3	34.7	2	37.6	2.3	35.5	1.6	31.5	1.6	26.3	3.3	18.6	4.4	12.8	2
2013	11.3	3.2	15	1.7	18.5	3.5	24	2.4	27.6	3.5	33	2	34.9	1.4	34	1.4	30.4	3	22.7	3.2	17.6	2	10	3.1
2014	11	2.3	12.7	4.2	18.9	2.8	24.7	5	30.5	2.7	33.7	2.7	36	1.4	36.7	1.8	31.9	2.3	25	3.4	16.2	3	13.5	2.5
2015	10.8	2.8	13.6	2.5	17.9	2	23.1	3.7	30.6	2.6	34.2	1.4	37.6	2	37.1	1.8	34	2.3	27.4	3.7	16.5	2.2	10.6	2.7
2016	10.4	3	14.8	2.9	18.5	2.5	24.7	3.8	29.8	2.6	34.4	2.6	37	1.6	36.6	1.2	31	2.8	25.9	1	15.3	4.3	10	3.3
2017	9.4	1.6	10.4	3.8	17.7	2.3	23.4	3.3	30.5	2.2	34.6	2.7	38.7	1.6	38.1	1.7	33.9	2.1	25.4	2.3	18.5	4	13.7	3
2018	12	2.3	14.9	2.3	21	2.9	23	2.4	28.9	3.4	34.2	1.3	36.2	1.7	35.5	1.2	32.9	2	27	4.7	17	1.4	13.2	1.9
2019	11	3	12.9	1.6	15.9	2.5	20.7	2.7	30.4	3.7	35.9	1.7	35.8	1	36.6	1.5	32.6	1.4	27	4	17.6	2.9	13.2	1.6
2020	11	1.7	13	4	18.5	2.7	23.8	2.7	30.4	3.2	33.8	1.6	38.4	1.8	34.9	1.6	33.9	2	25.8	3	19	3.8	12.4	2.4

Table (2- B): The daily mean temperature for the daily period and the standard deviation of the Basra stations.

Basra	January		February		March		April		May		June		July		August		September		October		November		December	
	Mean	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**	**Mean**	**S.D**
2010	15.5	2.5	18	3.9	23.3	3.6	27.6	2.1	33.2	3.3	38.5	2.4	39.4	1	39.6	1.2	35.6	1.7	29.9	3	20.3	3	15	2.2
2011	13.2	2.6	14.6	3.8	19.6	3.2	26.6	2.3	34	3.3	37.6	2	39.3	2	38.5	1.6	35.2	1.3	27.5	3.4	17.3	3.6	12.5	2.4
2012	12.7	3	14.3	2.3	18.7	3.7	27.4	2.6	34.8	2	38	1.8	40.9	1.9	39	1.5	34.6	1.7	29.6	2.8	20.9	4.2	15.2	2.3
2013	13.6	3	17.3	1.9	21.6	3.4	26.9	2.2	31.2	3.7	36.4	2	38.7	1	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**
2014	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	**m**	38.5	1.2	35.4	2	29	3.2	19.4	3	16.3	2.3
2015	14.1	3	16.9	2.7	21.3	2	27.1	2	34.3	2.7	37.8	1.3	40	1.5	40.4	1	36.2	1.6	31.1	3.2	20.6	2.5	13.2	3.1
2016	13.2	3	17.1	2.8	21.9	1.5	27	3.9	34	2.3	38.2	2.4	40.3	1.6	40	1.2	35	3	28.5	1	19.3	3.4	14.2	3.3
2017	13.5	1.6	13.5	3.5	20.9	16	28	3.4	34.5	2.2	38.2	2.7	41.2	1	40.3	1.4	36.4	1.7	29.7	2	21	4.3	15.6	2.2
2018	14.5	2	18	2.4	24	2.9	26.3	2	33.2	3.8	38.9	1.5	40	1.3	39.3	1	37.6	1.8	30	4.9	20	2.4	16.3	2.8
2019	14.5	2.8	16.2	1.5	19	2.4	25.4	2.5	34.3	2.7	39.5	1.6	39	1.3	39	1.1	37.6	1.9	30.8	3.7	20.3	2.9	15.3	2
2020	14	1.8	16.2	4.2	20.5	1.8	**m**	M	33.4	2.9	38.2	1.2	41.2	1.2	38.4	1.8	36.5	1.4	28	3.3	21.8	3.5	14.8	2.4

4. Result and Discussion

In the daily treatment, a comparison is made between the original values and values of the general mean to find the percentage of the difference, shown in Tables (3,4) for Baghdad and Basra stations, respectively. In Table (4), there are some missing data in the weather records symbolized by **m.**

Table 3: Percentage difference for the daily treatment in Baghdad station

Baghdad	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
2010	26%	43%	71%	23%	39%	35%	10%	19%	13%	45%	27%	23%
2011	16%	71%	52%	30%	42%	43%	26%	13%	10%	55%	43%	26%
2012	26%	38%	58%	33%	26%	23%	39%	23%	7%	55%	67%	19%
2013	48%	14%	52%	37%	55%	27%	0%	6%	37%	42%	20%	45%
2014	39%	61%	39%	60%	48%	43%	6%	19%	23%	52%	40%	32%
2015	42%	32%	23%	50%	39%	7%	23%	19%	33%	39%	30%	29%
2016	39%	52%	35%	50%	29%	40%	10%	6%	37%	0%	53%	39%
2017	16%	57%	29%	47%	23%	30%	3%	16%	17%	32%	57%	32%
2018	29%	25%	32%	27%	68%	3%	10%	6%	23%	71%	7%	16%
2019	48%	11%	26%	43%	61%	17%	0%	6%	3%	55%	50%	16%
2020	55%	16%	42%	40%	55%	7%	16%	3%	13%	32%	57%	26%

Table 4: Percentage difference in the monthly treatment in Basra station

Basra	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
2010	35%	43%	61%	27%	35%	33%	0%	0%	10%	58%	43%	26%
2011	39%	71%	48%	20%	42%	20%	19%	10%	3%	52%	63%	23%
2012	29%	21%	52%	33%	29%	17%	23%	6%	13%	35%	67%	26%
2013	42%	29%	39%	23%	48%	30%	**m**	**m**	**m**	**M**	**m**	**m**
2014	**m**	**m**	**m**	**m**	**m**	**m**	0%	0%	27%	48%	47%	10%
2015	39%	32%	19%	47%	35%	7%	6%	0%	10%	48%	43%	35%
2016	35%	41%	6%	47%	32%	37%	6%	0%	37%	0%	47%	52%
2017	10%	57%	10%	57%	29%	40%	0%	0%	17%	23%	67%	29%
2018	19%	32%	35%	23%	45%	10%	0%	0%	20%	61%	40%	48%
2019	52%	11%	29%	27%	35%	17%	6%	0%	20%	39%	40%	16%
2020	19%	59%	14%	**m**	52%	4%	4%	10%	10%	55%	63%	32%

Knowing the percentage of the difference led to the percentage of congruence of the arithmetic mean through Figures (2-3) that show the range of the match between actual and mean values for each month for the daily period. The mean temperature was not taken for several years due to the specificity of each year in terms of heat and cold, so the mean was taken for each month of each year. The two Figures (2-3) showed that the highest congruence percentage occurs in summer (June, July, August, and September). The highest percentage of congruence (100%, 100%) was for Baghdad and Basra stations, respectively, and the lowest percentage of congruence (57%, 60%) was for Baghdad and Basra stations, respectively. In the remaining months, the highest percentage of congruence (100%, 100%) was for Baghdad and Basra stations, respectively, and the lowest percentage of congruence was recorded for the same mentioned stations (29%, 29%). Moreover, the highest percentage of congruence was recorded in summer due to the stability of temperature, and investigating the data, shows that most of the summers that recorded 100% of congruence were in stable weather conditions.

Moreover, the highest percentage of congruence in winter was noticed in October 2016 for Baghdad and Basra stations, where October had a close temperature and stable weather conditions. The lowest percentage of congruence was in February 2011, due to a cold air mass that affected the temperature, which reduced it to the general mean. The arithmetic mean is very sensitive to the presence of any temperature anomalies, such as the too far from most temperature values. As the anomaly increases, the accuracy of the mean reduces. Therefore, the accuracy of the rate increases in summer, while in other months, it decreases.

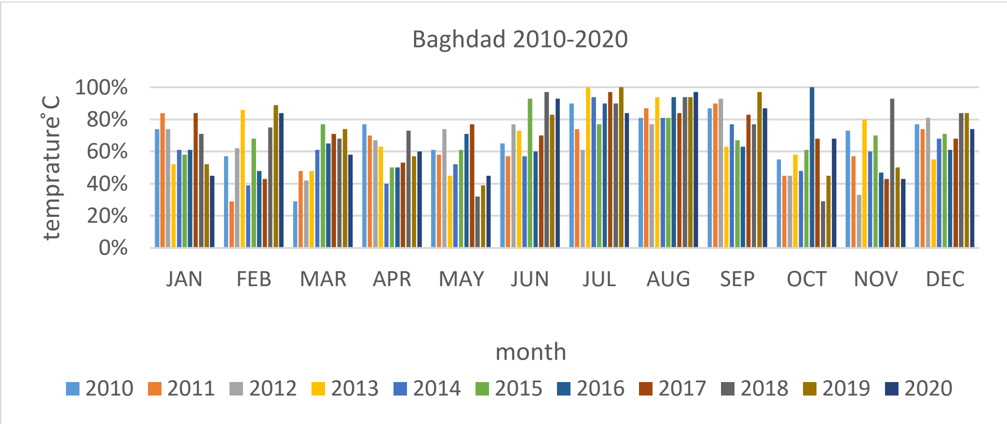

Figure 2: Percentage of mean-treatment-accuracy in Baghdad station

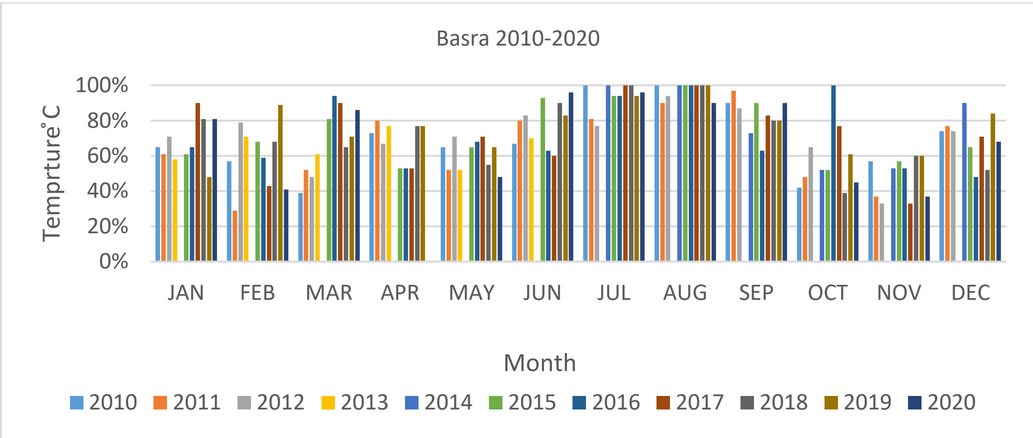

Figure 3: Percentage of mean-treatment-accuracy in Basra station

The second: is monthly; the general mean is taken for each month from (1980 - 2020). Through the comparison, noticing that there is a variance in the means for the entire period from the means of every ten years in most of the stations; so; the study period was divided into four periods (every ten years) for the stations (Baghdad, Hilla, Basra, Nasiriya, and Samawa), where the matching between the general mean value for each period and the original period were calculated based on the standard deviation values. The values obtained from the treatment are considered congruous with the actual values if the difference is less than (1.5) degrees. However, if the value was higher than (1.5) degrees, in that case, they are considered not close to the actual values, so values must be excluded and recalculated to get the percentage of difference, as shown in Tables (5, 6, 7, 8, and 9), displaying the mean values for each period, the standard deviation values, and the percentage of difference for all stations. These tables showed that every ten years, there is a mean that differed from the other periods, especially in the past ten years, where increased the mean value. Also, the value of the standard deviation increases in winter, spring, and autumn when the value of the standard deviation decreases as the progress towards summer (July, August, and September) so that the standard deviation becomes less in these months. Thus, the congruence of this method increases because the arithmetic is very sensitive to the presence of any abnormal values. This shows that the higher the standard deviation, the values increase above the general mean. The relationship between the standard deviation and the mean is inverse; as a result, when the period is divided partially, the treatment is better, and the mean is often closer to the actual values.

Table 5: Treatment Monthly in Baghdad station

month	estimator	year				The adopted mean estimator Through (1980-2019)
		1980-1989	1990-1999	2000-2009	2010-2019	
Jan.	mean	9.3	9.7	9.0	11	9.8
	standard deviation	1.8	1.7	1.2	1.2	1.6
	difference ratio	30%	40%	13%	10%	32%
Feb.	mean	11.6	11.8	13.0	13.4	12.6
	standard deviation	2.0	1.6	1.1	1.5	1.7
	difference ratio	40%	40%	25%	30%	39%
Mar.	mean	16.1	16.1	18.3	18.0	17.3
	standard deviation	1.2	1.6	1.5	1.7	1.8
	difference ratio	20%	30%	25%	30%	32%
Apr.	mean	22.7	22.7	23.6	23.7	23.3
	standard deviation	1.3	1.4	1.3	1.2	1.4
	difference ratio	10%	20%	38%	10%	29%
May.	mean	28.3	29.1	29.5	29.9	29.4
	standard deviation	1.3	1.1	1.1	1.0	1.2
	difference ratio	20%	30%	13%	10%	16%
Jun.	mean	32.4	32.5	33.7	34.4	33.4
	standard deviation	0.5	1.0	1.0	0.8	1.2
	difference ratio	0%	10%	13%	10%	13%
Jul.	mean	34.9	34.8	35.8	36.8	35.8
	standard deviation	0.7	0.7	0.9	1.1	1.2
	difference ratio	0%	0%	13%	20%	20%
Aug.	mean	33.9	34.2	35.3	36.4	35.2
	standard deviation	1.4	1.4	1.0	1.2	1.6
	difference ratio	30%	30%	13%	20%	37%
Sep.	mean	30.3	30.2	31.2	32.2	31.1
	standard deviation	1.0	0.7	0.9	1.3	1.3
	difference ratio	10%	0%	0%	30%	26%
Oct.	mean	23.8	24.5	25.3	25.8	25.1
	standard deviation	1.3	1.0	1.3	1.6	1.5
	difference ratio	30%	10%	13%	30%	34%
Nov.	mean	16.2	16.7	15.8	17.2	16.6
	standard deviation	1.9	1.3	0.8	1.7	1.6
	difference ratio	50%	10%	0%	30%	37%
Dec.	mean	10.7	11.4	11.6	12.2	11.7
	standard deviation	1.5	1.7	1.8	1.6	1.7
	difference ratio	40%	50%	38%	30%	45%

Table 6: Treatment Monthly in Hilla station

month	Estimators	year				The adopted mean estimator Through (1980-2019)
		1980-1989	1990-1999	2000-2009	2010-2019	
Jan.	mean	9.7	10.2	9.9	11.3	10.4
	standard deviation	1.8	1.8	1.4	1.0	1.6
	difference ratio	50%	33%	22%	10%	29%
Feb.	mean	12.4	12.4	13.6	13.8	13.1
	standard deviation	2.0	1.5	1.3	1.5	1.6
	difference ratio	42%	22%	33%	30%	37%
Mar.	mean	16.8	16.4	18.1	18.5	17.5
	standard deviation	1.0	1.6	1.5	1.7	1.7
	difference ratio	0%	22%	20%	30%	26%
Apr.	mean	23.0	22.8	23.6	23.9	23.4
	standard deviation	1.4	1.2	1.3	1.1	1.3
	difference ratio	29%	11%	20%	10%	25%
May.	mean	28.7	29.0	29.6	29.7	29.3
	standard deviation	1.0	1.3	0.6	1.0	1.1
	difference ratio	14%	30%	0%	10%	16%
Jun.	mean	32.1	33.0	33.5	33.8	33.2
	standard deviation	0.6	0.7	0.8	0.8	0.9
	difference ratio	0%	0%	0%	0%	11%
Jul.	mean	34.8	34.6	35.4	35.9	35.2
	standard deviation	0.9	0.9	1.4	0.9	1.1
	difference ratio	14%	0%	20%	0%	19%
Aug.	mean	34.1	34.0	35.1	35.4	34.7
	standard deviation	1.2	1.3	1.0	1.1	1.3
	difference ratio	29%	30%	10%	10%	24%
Sep.	mean	31.3	30.3	30.9	31.8	31.0
	standard deviation	0.8	0.5	1.0	1.1	1.0
	difference ratio	0%	0%	10%	10%	11%
Oct.	mean	25.1	24.8	25.5	25.7	25.3
	standard deviation	1.0	0.9	1.3	1.6	1.2
	difference ratio	14%	0%	20%	30%	22%
Nov.	mean	16.1	17.4	16.5	16.7	16.7
	standard deviation	1.2	1.2	0.9	1.7	1.3
	difference ratio	14%	20%	10%	30%	25%
Dec.	mean	11.8	12.2	11.6	12.1	11.9
	standard deviation	1.9	2.0	1.6	1.4	1.7
	difference ratio	50%	40%	40%	30%	42%

Table 7: Treatment Monthly in Basra station

month	Estimators	year				The adopted mean estimator Through (1980-2019)
		1980-1989	1990-1999	2000-2009	2010-2019	
Jan.	mean	12.1	12.6	12.0	13.8	12.6
	standard deviation	1.6	1.9	1.2	0.9	1.6
	difference ratio	40%	60%	13%	0%	35%
Feb.	mean	14.2	14.6	15.5	16.3	15.1
	standard deviation	1.8	1.4	1.1	1.7	1.7
	difference ratio	30%	30%	13%	56%	43%
Mar.	mean	18.9	19.0	20.6	21.0	19.8
	standard deviation	1.2	1.4	1.1	1.7	1.6
	difference ratio	30%	20%	13%	33%	24%
Apr.	mean	25.6	25.9	27.0	26.9	26.3
	standard deviation	0.9	1.2	1.6	0.8	1.3
	difference ratio	10%	20%	38%	0%	19%
May.	mean	31.6	32.4	33.7	34.1	32.8
	standard deviation	1.1	1.5	0.5	0.6	1.4
	difference ratio	10%	40%	0%	0%	31%
Jun.	mean	34.7	36.8	37.0	38.1	36.6
	standard deviation	0.4	1.1	0.4	0.9	1.5
	difference ratio	0%	10%	0%	0%	41%
Jul.	mean	37.0	37.8	38.7	39.8	38.3
	standard deviation	0.8	1.1	0.8	0.8	1.3
	difference ratio	0%	20%	0%	0%	32%
Aug.	mean	35.9	37.4	38.4	39.6	37.8
	standard deviation	1.4	1.5	0.9	0.6	1.8
	difference ratio	30%	30%	0%	0%	45%
Sep.	mean	33.0	33.8	34.1	35.7	34.1
	standard deviation	1.0	0.9	0.8	1.0	1.3
	difference ratio	10%	0%	11%	11%	21%
Oct.	mean	26.9	27.9	28.8	29.8	28.3
	standard deviation	1.3	1.0	0.9	0.8	1.4
	difference ratio	20%	10%	0%	0%	32%
Nov.	mean	19.4	20.4	19.6	20.0	19.8
	standard deviation	1.6	0.9	0.7	1.1	1.2
	difference ratio	20%	10%	0%	22%	18%
Dec.	mean	13.5	14.5	14.0	14.8	14.2
	standard deviation	1.4	1.9	2.1	1.4	1.7
	difference ratio	20%	30%	56%	22%	42%

Table 8: Treatment Monthly in Nasiriya station

month	Estimators	year				The adopted mean estimator Through (1980-2019)
		1980-1989	1990-1999	2000-2009	2010-2019	
Jan.	mean	11.64	12.12	11.41	12.99	12.06
	standard deviation	1.4	1.9	1.2	1.0	1.5
	difference ratio	30%	40%	11%	10%	28%
Feb.	mean	13.89	14.1	15.13	15.89	14.74
	standard deviation	1.9	1.5	0.8	1.6	1.7
	difference ratio	40%	40%	0%	40%	36%
Mar.	mean	18.52	18.79	20.69	21.55	19.87
	standard deviation	1.1	1.8	1.3	1.5	1.9
	difference ratio	20%	30%	11%	20%	36%
Apr.	mean	25.19	25.4	26.24	25.98	25.69
	standard deviation	1.0	1.1	1.6	1.2	1.2
	difference ratio	10%	10%	33%	30%	18%
May.	mean	31.04	31.93	32.69	32.90	32.13
	standard deviation	1.1	1.4	0.7	0.9	1.3
	difference ratio	10%	40%	0%	10%	21%
Jun.	mean	34.67	36.01	36.46	37.27	36.09
	standard deviation	0.5	0.8	0.8	0.9	1.2
	difference ratio	0%	0%	11%	20%	21%
Jul.	mean	37.04	37.26	38.39	39.17	37.96
	standard deviation	0.8	1.1	0.9	0.9	1.3
	difference ratio	0%	20%	11%	10%	23%
Aug.	mean	36.06	36.9	38.33	39.19	37.60
	standard deviation	1.4	1.5	0.9	0.9	1.7
	difference ratio	20%	30%	0%	10%	46%
Sep.	mean	33.19	33.51	33.74	35.44	33.98
	standard deviation	0.8	0.7	1.2	1.1	1.3
	difference ratio	0%	0%	11%	0%	21%
Oct.	mean	26.65	27.42	28.41	29.12	27.89
	standard deviation	1.2	1.0	1.2	1.4	1.5
	difference ratio	20%	10%	11%	10%	41%
Nov.	mean	18.99	19.64	18.91	19.46	19.26
	standard deviation	1.7	1.2	1.0	0.9	1.2
	difference ratio	40%	20%	0%	10%	18%
Dec.	mean	12.96	14.03	13.48	14.27	13.69
	standard deviation	1.5	1.8	2.0	1.4	1.7
	difference ratio	40%	40%	55%	30%	36%

Table 9: Treatment Monthly in Samawa station

month	Samawa	year				The adopted mean estimator Through (1980-2019)
		1980-1989	1990-1999	2000-2009	2010-2019	
Jan.	mean	11.2	11.3	10.0	12.7	11.4
	standard deviation	1.6	1.6	1.1	0.9	1.6
	difference ratio	30%	30%	28%	10%	35%
Feb.	mean	13.5	13.2	13.8	15.0	13.9
	standard deviation	2.0	1.3	1.5	1.7	1.7
	difference ratio	50%	20%	42%	40%	39%
Mar.	mean	18.2	17.3	19.2	20.8	18.9
	standard deviation	1.2	1.5	2.3	2.0	2.1
	difference ratio	10%	20%	42%	40%	46%
Apr.	mean	25.1	24.4	25.4	25.6	25.1
	standard deviation	1.1	0.9	1.5	1.3	1.2
	difference ratio	10%	10%	50%	30%	19%
May.	mean	30.5	31.0	32.0	32.3	31.4
	standard deviation	0.9	1.2	0.9	0.7	1.2
	difference ratio	0%	37%	14%	0%	23%
Jun.	mean	34.2	34.9	35.5	36.4	35.2
	standard deviation	0.4	0.7	0.8	0.8	1.1
	difference ratio	0%	0%	0%	0%	8%
Jul.	mean	36.5	36.1	37.0	38.1	37.0
	standard deviation	0.7	1.2	1.2	1.0	1.3
	difference ratio	0%	20%	14%	10%	28%
Aug.	mean	35.6	35.4	37.2	38.1	36.5
	standard deviation	1.4	1.6	1.1	1.0	1.7
	difference ratio	20%	40%	28%	10%	35%
Sep.	mean	32.8	32.4	32.7	34.1	33.0
	standard deviation	1.4	0.8	0.7	1.0	1.2
	difference ratio	10%	0%	0%	0%	16%
Oct.	mean	26.2	26.4	27.0	27.9	26.8
	standard deviation	1.2	0.9	1.5	1.6	1.4
	difference ratio	10%	10%	25%	30%	29%
Nov.	mean	18.4	18.5	17.5	18.9	18.4
	standard deviation	1.8	1.3	1.1	0.9	1.4
	difference ratio	30%	20%	0%	0%	29%
Dec.	mean	12.7	13.4	12.9	14.0	13.3
	standard deviation	1.4	2.2	2.1	1.1	1.7
	difference ratio	20%	40%	43%	20%	32%

The percentage of congruence in the monthly treatment is shown in figures (4, 5, 6, 7, and 8), noticing that the percentage of congruence is low when the mean method is applied to the total period for all months. Therefore, the study period was divided into parts, where all periods have a high matching rate in summer (100%), while in the winter, the matching percentage decreases to reach some periods to 44%.

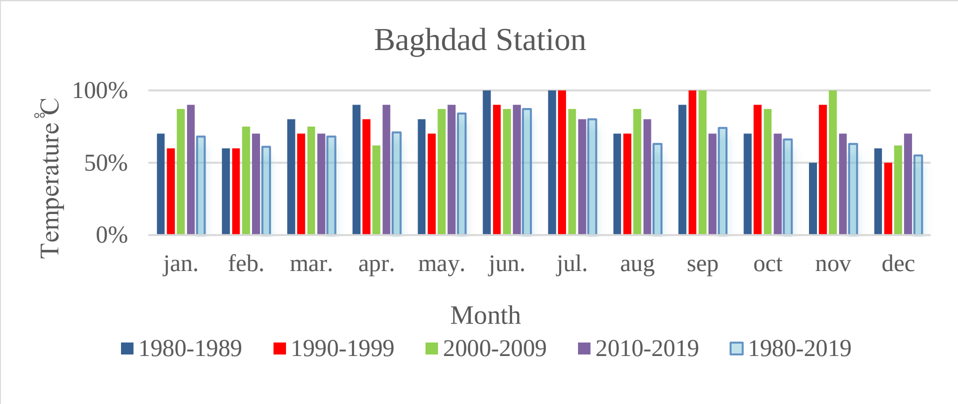

Figure 4: The accuracy ratio of the mean method for the Baghdad station

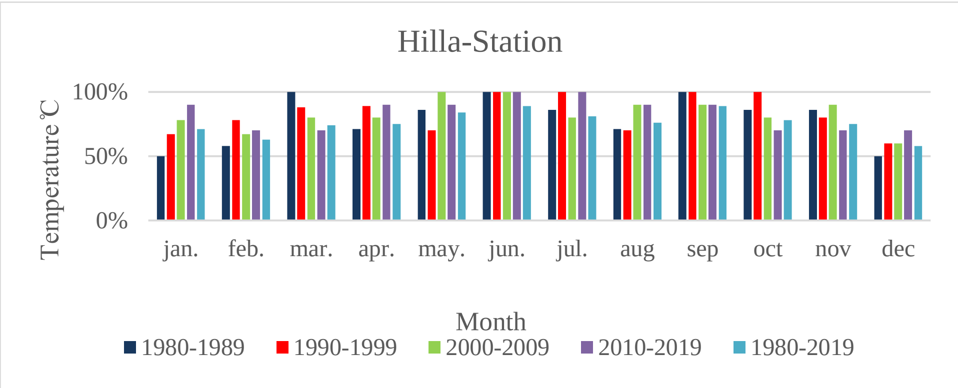

Figure 5: The accuracy ratio of the mean method for Hilla station

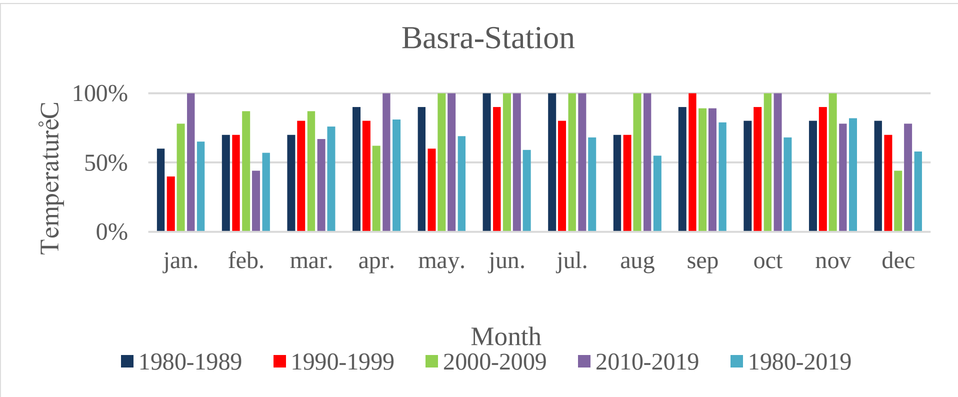

Figure 6: The accuracy ratio of the mean method for the Basra station

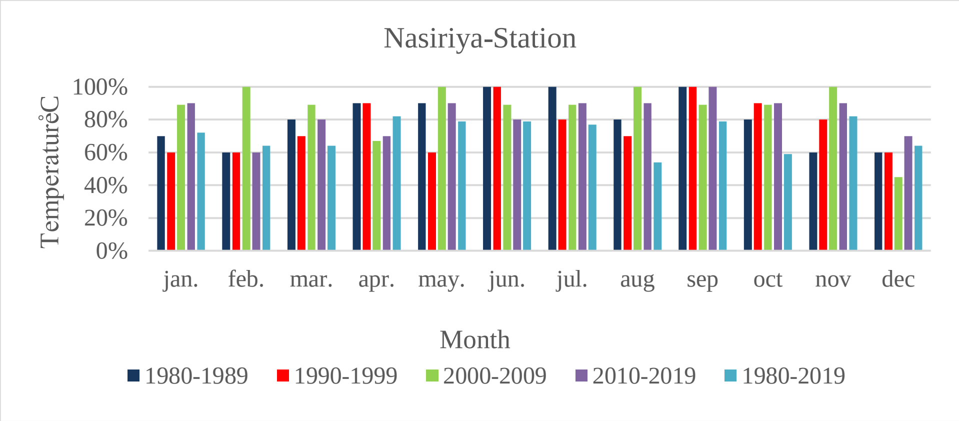

Figure 7: The accuracy ratio of the mean method for Nasiriya station

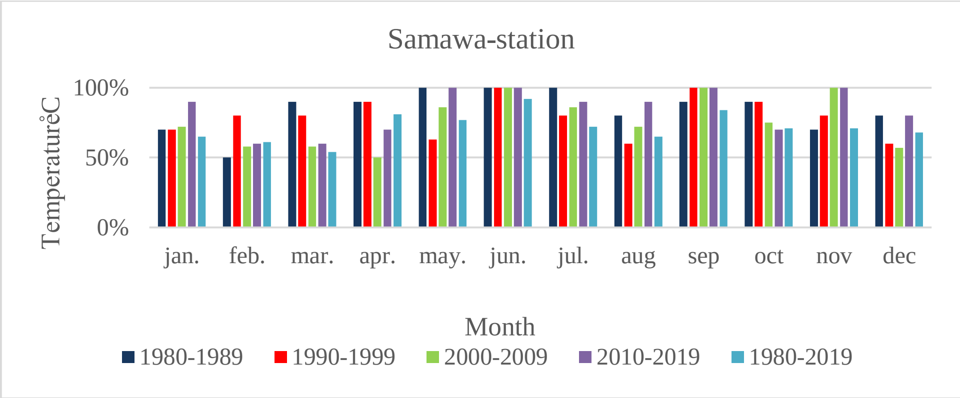

Figure 8: The accuracy ratio of the mean method for Samawa station

Conclusion

The data treatment was accomplished using the mean method, described as the easiest and fastest method to process data. The mean method was tested to find the missing temperature data for some stations in Iraq (central and southern regions). Using the method during summer produces very high accuracy, reaching 100%, unlike in winter, where the accuracy percentage was lower. In the monthly treatment, the results have shown that the study period prefers to be applied every ten years rather than over a more extended period. In the monthly treatment, the method is preferable and can produce a great result if there is one missing value (or ten years study period).

In comparison, in the case of more than one missing value (over three values), applying the method for a more extended period (20 years case study) is required. Eventually, five missing values require applying the method over a more extended period (30 years). In the daily treatment, the summer produces the highest accuracy rather than the other months. The congruence ratio reaches 100%, provided that the missing values do not exceed five values for the same month and must be non-consecutive for more than two days, which is not recommended for this method in the winter. Applying this method to the monthly treatment of missing data gives better accuracy than applying it to daily treatment.

References

[1] E. S. Ahmed and A. S. Hassan, "The impact of the extreme Air temperatures on the characteristics of Iraq weather," *Iraqi Journal of Science*, vol. 59, no.2C, pp. 1139-1145, 2018.

[2] A. Bunker, J. Wildenhain, A. Vandenbergh, N. Henschke, J. Rocklöv, S. Hajat and R. Sauerborn, "Effects of air temperature on climate-sensitive mortality and morbidity outcomes in the elderly; a systematic review and meta-analysis of epidemiological evidence," *EBioMedicine*, vol. 6, pp. 258-268, 2016.

[3] A. P. Kirilenko, T. Molodtsova, and S. O. Stepchenkova, "People as sensors: Mass media and local temperature influence climate change discussion on Twitter," *Global Environmental Change*, vol. 30, pp. 92-100, 2015.

[4] Y. Brugnara, J. Flückiger, and S. Brönnimann, "Instruments, procedures, processing, and analyses," 2020.

[5] E. C. Kent, N. A. Rayner, D. I. Berry, R. Eastman, V. G. Grigorieva, B. Huang J.J. Kennedy, S.R. Smith, and K.M. Willett, "Observing requirements for long-term climate records at the ocean surface," *Frontiers in Marine Science*, vol. 6, p. 441, 2019.

[6] G. Manogaran, D. Lopez, and N. Chilamkurti, "In-Mapper combiner-based MapReduce algorithm for processing of big climate data," *Future Generation Computer Systems*, vol. 86, pp. 433-445, 2018.

[7] I. Gad and B. Manjunatha, "Performance evaluation of predictive models for missing data imputation in weather data," in 2017 *International Conference on Advances in Computing, Communications and Informatics* (ICACCI), 2017, pp. 1327-1334.

[8] R. Arsenault and F. Brissette, "Determining the optimal spatial distribution of weather station networks for hydrological modeling purposes using RCM datasets: An experimental approach," *Journal of Hydrometeorology*, vol. 15 ,no.1, pp. 517-526, 2014.

[9] E. R. Buhi, P. Goodson, and T. B. Neilands, "Out of sight, not out of mind: Strategies for handling missing data," *American journal of health behavior*, vol. 32,no.1, pp. 83-92, 2008.

[10] S. T. Nassir, A. B. Khamees, and W. T. Mousa, "Estimation the Missing Data of Meteorological Variables In Different Iraqi Cities By using ARIMA Model," *Iraqi Journal of Science*, vol. 59, no. 2A, pp. 792-801, 2018.

[11] J. Rahimi, A. Khalili, and J. Bazr Afshan, "Evaluation of different missing data reconstruction methods for daily minimum temperature in elevated stations of Iran: comparison with new proposed approach," *Iranian Journal of Soil and Water Research*, vol. 48 ,no.2, pp. 231-239, 2017.

[12] M. Popolizio, A. Amato, T. Politi, R. Calienno, and V. Di Lecce, "Missing data imputation in meteorological datasets with the GAIN method," in 2021 *IEEE International Workshop on Metrology for Industry 4.0 & IoT* (MetroInd4. 0&IoT), 2021, pp. 556-560.

[13] H. T. Majeed, W. G. Nassif, and Y. Q. Tawfeek, "Effects of meteorological parameters on surface concentration of carbon monoxide over Iraq," *J. Green Eng*, vol. 10 ,no.9, pp. 5927-5940, 2020.

[14] M. Al-Mukhtar and M. Qasim, "Future predictions of precipitation and temperature in Iraq using the statistical downscaling model," *Arabian journal of geosciences*, vol. 12 ,no.2, pp. 1-16, 2019.

[15] WMO, "Guide to Instruments and Methods of Observation," ed: *World Meteorological Organization Geneva*, 2018.

[16] N. M. Noor, M. M. Al Bakri Abdullah, A. S. Yahaya, and N. A. Ramli, "Comparison of linear interpolation method and mean method to replace the missing values in environmental data set," in *Materials Science Forum*, 2015, pp. 278-281.