Improving the estimation of the true mean monthly and true mean annual air temperatures in Greece

N. K. Sakellariou and H. D. Kambezidis*
Atmospheric Research Team, Institute for Environmental Research & Sustainable Development, National Observatory of Athens, Greece

*Correspondence to:
H. D. Kambezidis,
Atmospheric Research Team,
Institute for Environmental Research & Sustainable Development, National Observatory of Athens, GR-11810 Athens, Greece.
E-mail: harry@noa.gr

Abstract
The true mean monthly/annual air-temperature estimations are usually based on the monthly averages of meteorological observations performed at standard hours. The performance of this combination varies as it depends on local climate of the particular meteorological station; it consequently allows for small/large deviations from its exact value. This study examines the possible deviations in Greece and suggests ways of improving the derived estimates. Two cases are considered. (1) An area has one thermograph and a number of thermometric stations. (2) An area has several thermographs and many thermometric stations. Solutions for minimizing the estimation error in both cases are provided.

Keywords: temperature estimation; true average; climate normals; Greece

1. Introduction
At meteorological stations equipped with thermographs, continuous air-temperature recordings are available. From such observations, 24-h air-temperature values can be deduced. Their arithmetic mean is defined as the true mean temperature or the true daily mean (Conrad and Pollak, 1950; Weiss and Hays, 2005; Dall’Amico and Hornsteiner, 2006), while the monthly average of the true daily means yields the true monthly mean.

When only a few (usually three) daily air-temperature observations are available, a linear combination of the monthly average temperatures at the standard hours of the observations is used in order to obtain the closest possible approximation to the true monthly mean. These combinations depend on the climate of the station; Conrad and Pollak (1950) have listed the formulas that obtain best results for the greater part of Europe as well as for tropical and subtropical climates.

All the formulas can give an accurate answer, provided that there is no daily temperature variation. This implies that the formulas tend to express the average daily air-temperature variation of the stations that correspond to the climate; they consequently fail when the daily air-temperature variation at a particular station strongly deviates from its assumed pattern. This failure may yield miscalculations concerning the heating demands of buildings (e.g. Kaufmann et al., 2013) as well as predictions of crop production (e.g. Rosenzweig and Parry, 1994; Olesen and Bindi, 2002).

The formula that has been suggested by the aforementioned authors as very advantageous for the greater part of Europe is:

$$\overline{T} = \frac{(T_7 + T_{14} + 2T_{21})}{4}$$  \hspace{1cm} (1)

where $\overline{T}$ is the approximate true average monthly air temperature and $T_7$, $T_{14}$, and $T_{21}$ are the monthly mean air temperatures at the 0700, 1400, and 2100 of observation, respectively. The hours refer to mean local time, UTC+2.

A modified version of this formula, with $T_8$ replacing $T_7$, has been used by Aeginitis (1907) and Mariolopoulos (1938) for the estimation of the true mean monthly temperature in Greece. The validity of the modified formula has been tested with the help of the thermographic recordings taken by a Richard-type thermograph at the meteorological station of the National Observatory of Athens (hereafter NOA) from 1894 to 1903 (Aeginitis, 1907). The error of the approximation was found to be small at NOA (Aeginitis, 1907) and of the order of 0.1°C in most cases (Mariolopoulos, 1938). The formula:

$$\overline{T} = \frac{(T_8 + T_{14} + 2T_{21})}{4}$$  \hspace{1cm} (2)

was used all over Greece for the estimation of the true monthly air temperature under the implied assumption that it performs equally well all over the country.

The latter constitutes a rather optimistic point of view, considering the climatic diversification of Greece. The aim of this study is, therefore, to investigate the possible errors that may arise as a result of the discriminatory application of this formula and to suggest ways to reduce these errors.
of improving the true mean monthly temperature estimates.

It should be noted that Equation (2) is still in use by the Hellenic National Meteorological Service (HNMS) in a slightly different form: the observations taken at 2100 local time have been replaced by those at 2000 from 1930 onwards.

Two different cases are considered here for this problem. First, when only one thermograph operates in a region but thermometric observations are available at a number of locations. This case resembles the situation in Greece from 1894 to 1930. Second, thermographic observations are available at a limited number of sites and thermometric ones exist over a denser network of meteorological stations. This resembles the situation in Greece from 1950 onwards. Solutions to the problem are suggested in this study for both cases and the results are inter-compared. The objective of this study is to provide corrections to the pre-1930 temperature measurements when thermometric stations mainly existed.

2. Case 1. True mean air-temperature estimations with one thermograph

It should be mentioned here that a typical temperature reading difference between a thermometer and a thermograph may be around 0.5–1°C (Srivastava, 2008). At NOA’s meteorological station a 2-year analysis of the temperature difference between the dry-bulb thermometer and the bimetallic thermograph has shown this to be at 0.5°C on average, with the thermograph giving lower temperature values.

In order to improve the air-temperature estimates deduced by Equation (2), it is required that the following are secured:

1. The average hourly and daily values for every month and for a number of years should be available from the operating thermograph; its location is called principal location hereafter.
2. The average hourly values for each month should be available at a number of meteorological stations, where air-temperature observations are taken at standard hours every day; these sites are called secondary locations hereafter.
3. A procedure that takes into account the aforementioned thermographic and thermometric data and produces improved true mean monthly estimates at the secondary locations should also be developed.
4. Supplementary thermographic data at the secondary locations should be available, in order to validate the new estimates.

NOA has been chosen as the primary location; its registered thermographic observations taken in the period 1916–1930 have been used in this study. The secondary locations considered belong to the HNMS network and there has been given attention to be distributed evenly across Greece; their data are used in this study.

| No. | HNMS station                      | Years  |
|-----|-----------------------------------|--------|
| 1   | Agrinio                           | 1992, 1993 |
| 2   | Arta                              | 1992, 1993 |
| 3   | HNMS headquarters at Elliniko (Athens area) | 1993 |
| 4   | New Philadelphia (Athens area)     | 1992 |
| 5   | Ierapetra                         | 1992, 1993 |
| 6   | Ioannina                          | 1992, 1993 |
| 7   | Kastoria                          | 1992, 1993 |
| 8   | Kerkyra (Corfu)                   | 1992, 1993 |
| 9   | Kozani                            | 1992, 1993 |
| 10  | Larissa                           | 1992, 1993 |
| 11  | Trikala                           | 1992, 1993 |

Another requirement of the study is the availability of thermographic recordings for validation purposes, in parallel to the availability of thermometric observations at the same stations, even for a small number of years. Table 1 and Figure 1 give a full list of the meteorological stations and the corresponding years of measurements used in this study. These stations have thermographic recordings. The low number of years for the HNMS data used in this study depends upon their verified reliability.

The proposed algorithm takes into account the true mean monthly air-temperature value \( T \) and its monthly averages at the standard times of observations \( T_{8}, T_{14}, T_{21} \) at the principal location. These standard times were in use in Greece during the period of 1894–1930.

At a secondary location, X, the corresponding true mean monthly air temperature, \( T_{X} \), is assumed equal to \( T \), when the corresponding monthly averages at the standard hours of observations are equal to those at the principal location. If one of those observations differs and the remaining two are equal, it is assumed that the difference decreases linearly with time and becomes zero at the other two observational hours. If,
for example, $T_{X14}$ is greater than $T_{14}$, their difference is:

$$D_{14} = T_{X14} - T_{14} \quad (3)$$

which decreases linearly to zero at 0800 and 2100; then $T_X$ becomes greater than $T$:

$$T_X - T = D_{14} \frac{6.5}{24} \quad \text{or}$$

$$T_X = T + D_{14} \frac{6.5}{24} \quad (3.1)$$

Similarly:

$$T_X = T + D_{8} \frac{8.5}{24} \quad (3.2)$$

and:

$$T_X = T + D_{21} \frac{9}{24} \quad (3.3)$$

If all monthly averages at the standard hours of observations differ, then, by combining Equations (3.1)–(3.3), we deduce:

$$T_X = T + \frac{8.5D_8 + 6.5D_{14} + 9D_{21}}{24} \quad (4)$$

Equation (4) is the mathematical expression of the proposed algorithm; it implies that maximum differences of the averaged hourly air temperatures between station X and NOAA occur at the standard hours of observations. This is rather optimistic, but as the actual daily course of the air temperature at station X in unknown, Equation (4) is expected to yield better approximation estimates of the true monthly means at station X than that of Equation (1). In practice, Equation (2) can be regarded as the first-order approximation to the true monthly mean at station X, and Equation (4) can serve as the second-order approximation.

### 3. Validation of the algorithm

Equations (2) and (4) were applied to the data from all stations and years listed in Table 1 and Figure 1. For every month, year, and station, the differences between the actual true mean air temperatures deduced from thermographic recordings were compared with the estimations derived from Equations (2) and (4).

The obtained differences, averaged over all stations and years of observations, are listed in Table 2; it is noticed that from late spring until autumn, Equation (2) overestimates the true mean monthly air temperatures far more than previously anticipated. The averaged differences exceed 0.8°C in June; occasionally, the difference can become higher than 1.0°C at inland stations. Application of Equation (4) alleviates this problem; it corrects more than 70% of the occurring error from March until September. During these months (except for March and September), the results from Equation (4) are found to produce better estimates of the true mean monthly temperatures than those received from Equation (2) at the significance level of 99%. During the remaining months, Equation (4) performs worse than Equation (2), with an exception for February and December. Equation (2) gives moderate overestimated true mean monthly temperatures, while Equation (4) has a tendency toward considerable underestimation. This behavior of Equation (4) can partly be attributed to the breakdown of the hypothesis for a linear decrease of the observed differences $D_8$, $D_{14}$, and $D_{21}$ with time. The shorter length of the day may modify and complicate the daily course of the temperature during the months of October through February.

True mean annual temperatures can be deduced from true mean monthly values. As Equation (4) provides both overestimated values for some months and underestimated for others, it approaches, on the average, the true annual mean with a very good accuracy (0.011), while Equation (2) gives an average overestimation by almost 0.385°C. This overestimation varies from year to year and, thus, considerable noise can be introduced into the time-series analysis that tries to estimate temperature-related climate change, when Equation (2) is used in order to deduce the mean annual temperature values.

### 4. Case 2. True mean air-temperature estimations with a number of thermographs

In this scenario, a limited number of thermograph-equipped stations exist (more than one principal location). The problem to be solved is the estimation of the true mean monthly air temperatures at a number of secondary locations, where only daily temperature observations at standard hours are available.

As more thermographs are now available, it is possible to express their true mean monthly $T$ as a function of the monthly temperature averages $T_8$, $T_{14}$, and $T_{21}$, taken at the corresponding standard observational hours with the help of linear multiple regression analysis. Consequently, 12 equations, one for each month, are

### Table 2. Differences between true mean monthly air-temperatures derived from thermographic recordings and corresponding estimations deduced from Equations (2) and (4).

| Month       | Difference from Equation (2) (°C) | Difference from Equation (4) (°C) |
|-------------|----------------------------------|----------------------------------|
| January     | 0.074                            | −0.580                           |
| February    | 0.167                            | −0.115                           |
| March       | 0.274                            | −0.072                           |
| April       | 0.504                            | 0.060                            |
| May         | 0.602                            | 0.116                            |
| June        | 0.846                            | 0.287                            |
| July        | 0.792                            | 0.162                            |
| August      | 0.657                            | 0.139                            |
| September   | 0.382                            | −0.042                           |
| October     | 0.130                            | −0.202                           |
| November    | 0.064                            | −0.110                           |
| December    | 0.098                            | 0.008                            |
| Year        | 0.385                            | 0.011                            |
The linear fit has features for January for all stations and years listed in Table 1. Such equations, derived from the observations made at the stations and years listed in Table 1, are presented in Table 3. Comparing the coefficients derived. Such equations, deduced from the observations made at the stations and years listed in Table 1, are presented in Table 3. Comparing the coefficients of these equations to the coefficients of Equation (2) (not shown here), it is realized that only the coefficient at noon is similar in all equations. This is another indication of the inadequacy of Equation (2).

Observed versus predicted values based on the aforementioned equations and for the months of January and July are presented in Figures 2 and 3, respectively, in order to depict the accuracy of the fit. From these equations and using the temperature observations at the same observational times at the secondary locations, the true monthly means at these locations can be estimated. It is implied that the daily temperature course at the secondary locations is more or less similar to the average daily temperature variation of the stations at the primary locations. Otherwise, the estimated values are of poor quality.

It should be pointed out although that the procedure of Section 2 can also be employed for the true mean monthly temperature estimations at the secondary locations. This entails retaining only one primary location and omitting the thermographic recordings of the others. Consequently, the question about the most promising procedure arises.

To resolve this, another data set made available by HNMS stations with simultaneous thermographic recordings and temperature observations was employed. The stations and years of observations used in the subsequent analysis are listed in Table 4 and Figure 4.

True monthly means were estimated for every particular month, for all years and stations employed with the help of Equations (2) and (4), as well as with the help of equations listed in Table 3. Differences between observed and estimated true monthly means for every particular month and estimation procedure, averaged over all stations and years of observations are listed in Table 5, along with their seasonal and annual values.

Comparing Tables 2 and 5, it is observed that Equation (4) yields better estimates than Equation (2) from April to September. During the same period, the regression analysis equations yield more or less results of equal quality to those obtained through Equation (4). The results obtained from the regression analysis equations tend to be rather symmetrically distributed around the true monthly averages, while those deduced from Equations (2) and (4) tend toward overestimated values. It is, therefore, expected that the seasonal and annual averages obtained through the regression

| Month    | Equation                                                                 |
|----------|---------------------------------------------------------------------------|
| January  | $T = -0.58 + 0.32T_8 + 0.35T_{14} + 0.35T_{21}$                           |
| February | $T = -0.95 + 0.26T_8 + 0.34T_{14} + 0.46T_{21}$                         |
| March    | $T = -0.16 + 0.28T_8 + 0.27T_{14} + 0.45T_{21}$                         |
| April    | $T = -0.01 + 0.32T_8 + 0.25T_{14} + 0.41T_{21}$                         |
| May      | $T = -0.53 + 0.34T_8 + 0.26T_{14} + 0.40T_{21}$                         |
| June     | $T = 0.02 + 0.34T_8 + 0.23T_{14} + 0.40T_{21}$                          |
| July     | $T = 0.67 + 0.42T_8 + 0.17T_{14} + 0.38T_{21}$                          |
| August   | $T = 0.87 + 0.37T_8 + 0.25T_{14} + 0.25T_{21}$                          |
| September| $T = -0.14 + 0.31T_8 + 0.23T_{14} + 0.47T_{21}$                        |
| October  | $T = 0.26 + 0.35T_8 + 0.32T_{14} + 0.31T_{21}$                         |
| November | $T = 0.07 + 0.40T_8 + 0.31T_{14} + 0.29T_{21}$                         |
| December | $T = -0.32 + 0.34T_8 + 0.31T_{14} + 0.37T_{21}$                        |

Figure 2. Predicted (corrected) versus observed air temperatures for January for all stations and years listed in Table 1. The linear fit has $R^2 = 0.997$ and equation $y = -0.062 + 1.024x$ ($x =$ observed temperature, $y =$ predicted temperature).

Figure 3. As in Figure 2, but for July. The linear fit has $R^2 = 0.983$ and equation $y = 0.596 + 0.969x$ ($x =$ observed temperature, $y =$ predicted temperature).

Table 3. Multiple regression analysis equations derived from the observations made at the stations and years listed in Table 1.

| Month    | Equation                                                                 |
|----------|---------------------------------------------------------------------------|
| January  | $y = bx + a$                                                              |
| February | $y = bx + a$                                                              |
| March    | $y = bx + a$                                                              |
| April    | $y = bx + a$                                                              |
| May      | $y = bx + a$                                                              |
| June     | $y = bx + a$                                                              |
| July     | $y = bx + a$                                                              |
| August   | $y = bx + a$                                                              |
| September| $y = bx + a$                                                              |
| October  | $y = bx + a$                                                              |
| November | $y = bx + a$                                                              |
| December | $y = bx + a$                                                              |

Table 4. HNMS stations and years of observations employed for the inter-comparison of the results obtained from Equations (2) and (4) and those from the regression analysis of Table 3.

| No. | HNMS station | Years         |
|-----|--------------|---------------|
| 1   | Kalamata     | 1991, 1992, 1993 |
| 2   | Argostolion  | 1991, 1992    |
| 3   | Aliartos     | 1991, 1992, 1993 |
| 4   | Alexandroupolis | 1991, 1992, 1993 |
| 5   | Argos        | 1991, 1992, 1993 |
| 6   | Araxos       | 1991, 1992, 1993 |
| 7   | Aghialos     | 1991, 1992, 1993 |
| 8   | Zakynthos (Zante) | 1991, 1993   |
| 9   | Iraklion     | 1992, 1993, 1994 |

*The location of the stations is shown in Figure 4.
analogy equations should yield better estimates than those obtained through Equations (2) and (4). The annual averages listed in Table 5 demonstrate this. Equation (2) can be retained for winter and late autumn months, as it performs more or less equally well with the other two methods.

5. Possible effect on climatic temperatures of Greece

The first tabulation of the mean monthly temperatures, based on observations taken over a long-term period from many meteorological stations operating all over Greece, appears for the first time in the ‘Climate of Greece’ by Mariolopoulos (1938). Mean temperatures were derived with the help of Equation (2) and it was assumed that they represent the true means with great accuracy. The climatic period was 1900–1929, but as many of the stations started their operation later than 1900 (about 1915) the means of the period 1915–1929 had to be adjusted to yield the 1900–1929 values. The adjustment was based on the already derived mean values from the stations with full records.

To assess the possible error introduced in the true mean climatic estimations, observations taken at a number of stations listed in Table 6 and for periods of observations varying from 1910–1929 and 1915–1929 were used in order to obtain averages at the standard hours of observations for every month. Their data were taken from the NOA archives. It should be mentioned that before 1930, NOA was responsible for all meteorological observations in Greece.

From these averages, the monthly means were estimated with the help of Equation (2) as well as with the help of the regression analysis equations. The latter are considered representing the true means; the differences between their estimates and the corresponding estimates from Equation (2) are considered representing the error in the climatic true mean air-temperature estimates.

Comparing Tables 5 and 6, it is noticed that the general characteristics of the overestimated values obtained from Equation (2) are exhibited in both tables, implying that the overestimations of Equation (2) concern not only individual years but also the large observational periods as well. Maximum errors are found at stations located in northern Greece and away from the sea, while minimum ones occur at stations located in the southern part of the country and close to the sea, as can be seen in Figure 5, where the locations of the stations are shown. The annual error equals approximately 50% of the error of the summer season in both tables. As mentioned in Section 1, the meteorological observations at 2100 local time were replaced by those at 2000 local
6. Conclusions

From the present study, it is evident that Equation (2) largely overestimates the true mean air temperatures, with the exception of winter and late autumn months. Overestimations occur for individual months, years, or climatic periods. They reach maximum values in the summer months and for stations located in northern Greece and away from the sea. The opposite holds for stations located in the southern part of the country and close to the sea. Two solutions for the problem have been suggested.

The first is applied when only one thermograph is operating in an area and its recordings should be used along with their air-temperature observations taken at other locations, in order to obtain the true mean air temperature at these locations. Equation (4) is employed for the solution to the problem and corrects more than 50% of the overestimation error in spring, summer, and the annual true mean air temperatures.

The second solution is applied when more than one thermographs is operating in an area. Multiple regression analysis is employed and the true mean air temperatures for these stations are expressed in the form of equations having the true mean value as dependent variable and the mean monthly averages at the standard observation times as independent ones (see equations in Table 3). These equations yield the true mean values at locations, where only air-temperature observations at the standard observation hours are available.

The two solutions yield comparable results for individual spring and summer months, while the second solution yields better seasonal and annual estimates. The annual averages obtained through the second solution are more accurate than those obtained from Equation (2). Therefore, the time series of the mean annual air-temperature values derived from Equation (2) may contain noise, thus obscuring the detection of any climate change signal.

Acknowledgements

Thanks to the HNMS are expressed for providing the air-temperature data (thermometric and thermographic) for the mentioned stations and years in the study. The authors also declare that they have no conflict of interest.

References

Aeginitis D. 1907. The Climate of Greece. Part A: The Climate of Athens.: P. Sakkelariou Editions, Athens; 229 pp.
Conrad V, Pollak WL. 1950. Methods in Climatology, 2nd ed. Harvard University Press: Cambridge, Massachusetts; 154–160.
Dall’Amico M, Hornsteiner M. 2006. A simple method for estimating daily and monthly mean temperatures from daily minima and maxima. *International Journal of Climatology* 26(13): 1929–1936.
Kauffmann RK, Gopal S, Tang X, Raciti SM, Lyons PE, Geron N, Craig F. 2013. Revisiting the weather effect on energy consumption: implications for the impact of climate change. *Energy Policy* 62: 1377–1384.
Mariopoulos H. 1938. The Climate of Greece. Papaspyrou Press, Athens; 19–21.
Olesen JE, Bindi M. 2002. Consequences of climate change for European agriculture productivity, land use and policy. *European Journal of Agronomy* 16(4): 239–262.
Rosenzweig C, Parry ML. 1994. Potential impact of climate change on world food supply. *Nature* 367: 133–138.
Srivastava GP. 2008. *Surface Meteorological Instruments and Measurement Practices*. Atlantic Publishers and Distributors Ltd, New Delhi, 978-81-269-0968-1.
Weiss A, Hays CJ. 2005. Calculating daily mean air temperatures by different methods: implications from a non-linear algorithm. *Agricultural and Forest Meteorology* 128: 57–65.