

Degree-Day Formulations and Application in Turkey

MIKDAT KADIOĞLU AND ZEKAI ŞEN

Meteorology Department, Istanbul Technical University, Istanbul, Turkey

(Manuscript received 3 September 1997, in final form 3 August 1998)

ABSTRACT

Degree-days are fundamental design parameters in many application fields such as power generation and consumption, agriculture, architecture, snow melt estimation, environmental energy planning, population siting, and military domains. Depending on temperature fluctuations, the degree-day statistics at any location show local and seasonal variations. Among these parameters the average degree-day durations for cooling and heating periods, degree-day sums, and their maximums play a significant role in practical applications. In the body of literature to date the average degree-day durations have been analytically studied most often for independent processes. In this paper, however, degree-day sums in addition to durations are considered as important design variables with analytical formulation for dependent processes on the basis of the first-order Markov process. The application of the methodologies developed are presented for five temperature measurement stations scattered throughout Turkey within different climate regions.

1. Introduction

Temperature records are among the most common meteorological parameters after precipitation in describing the climate of a region. In addition, the temperature records in a region are primary indicators of the human and plant activities that take place over time. Agricultural practices, architectural designs, power generation, and effects of snowmelt, freezing, and icing on the transportation systems are all related to the temperature fluctuations. Plant growth, flowering and harvesting dates, additional loads on electrical power for heating or cooling in the big cities are all related to relative values of expected daily temperature with respect to a design level, that is, threshold. This threshold level depends on the specific purpose. For example, plant growth may not occur at temperatures less than 7°C, and the comfort level of human beings begins at 18°C. The intensity of the growing period, as opposed to its length, is frequently represented by the accumulation of temperature units above the plant growth or human comfort threshold. The basic unit is the degree-day, which takes into account the amount by which the daily mean temperature exceeds the stated minimum. Although recognized as an important climate variable, daily temperature time series are not subjected to detailed study so far as being truncated above or below certain levels. There is a need

for a relationship between the temperature levels and degree-day statistics in order to determine design temperature value given human comfort or plant growth temperatures. Provided that temperature records are available, the numerical calculation of degree-day statistics for any station is achieved easily. In practice, however, it is preferable to have empirical relationships between the degree-day and temperature statistics.

The first analytical approach to degree-day formulation was started by Thom (1952), who theoretically considered daily temperature to have a Gaussian independent normal distribution. Standard statistical analysis was applied in order to obtain the relationship between the means of the daily temperature records and degree-days derived therefrom. Thom (1954) further explained the relationship between mean temperature and the mean degree-days for a truncation level (base level) of 65°F. He showed that this relationship is independent of the base value. Later, Thom (1966) developed equations for obtaining mean monthly degree-days above any base-level temperature by the standard deviation of monthly average temperatures. Quayle and Diaz (1980) have observed that heating degree-days are directly related to site-specific total energy and heating oil consumption for individual residences. They employed a simple regression technique to depict the relevant relationship between the heating degree-days and consumed fuel oil. In this manner, they concluded that accurate weather projections could yield accurate short-term energy demand projections. Guttman (1983) analyzed the variability of population-weighted seasonal heating degree-days for 48 states in the United States.

Corresponding author address: Dr. Mikdat Kadioglu, Department of Meteorology, Istanbul Technical University, Ucak ve Uzay Bilimleri Fakultesi, Maslak, 80626 Istanbul, Turkey.
E-mail: kadioglu@itu.edu.tr

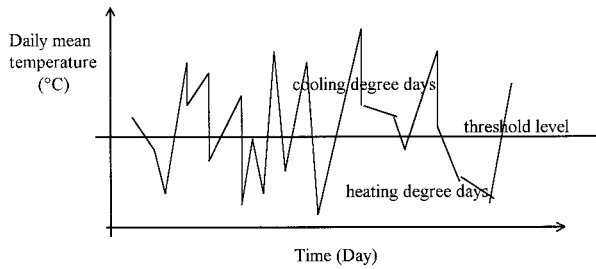


FIG. 1. Cooling and heating day definitions.

He concluded that the prediction of future heating-energy demand is dependent not only on historical averages but on the variability of the climate. Lehman (1987) examined the question of how very effective the skewness of the daily temperature variable is in estimating the mean value and variance of the degree-day variable at U.S. stations where the relationship is nonlinear. Later, Guttman and Lehman (1992) modeled the mean daily degree-hours by assuming a normal probability distribution for temperature. They proposed four different types of models and tried to assess their validity by considering the underlying normality and constant variance assumptions. They showed that mean daily degree-hours were found to be best expressed as a function of the standardized truncated normal distribution of the difference between the base temperature and hourly mean temperature.

This study analyzes various statistical properties of the temperature series truncated at a certain constant level. The temperature values greater than this truncation level give rise to various degree-day quantities, such as the durations and sums. These quantities are obtained with rather simple analytical derivations and applied to daily temperature data from various parts of Turkey.

2. Basic definitions

The degree-day definition requires the truncation of a temperature series above or below a chosen threshold, as shown in Fig. 1. The temperature series can be analyzed statistically in order to find the probability density function of the temperature at any station in addition to the low-order statistical parameters, such as the mean, standard deviation, skewness coefficient, etc. However, the truncation of such a series gives rise to significant new definitions, hence new variables, which have practical impacts in many engineering and human activities. The following definitions along with Fig. 1 provide information about the degree-day characteristics.

(a) Any temperature, T_i , at time instant i greater than the base level T_b is named as the cooling temperature. The cooling amount T_{ci} can be expressed as

$$T_{ci} = T_i - T_b \quad (\text{if } T_i > T_b). \quad (1)$$

Similarly, the heating amount T_{hi} is defined as

$$T_{hi} = T_b - T_i \quad (\text{if } T_b > T_i). \quad (2)$$

(b) An uninterrupted sequence of cooling (heating) amounts preceded and succeeded by at least one heating (cooling) amount is called the cooling (heating) duration. For instance, a cooling duration of length 6 is defined notationally as $T_{wi}, T_{ci+1}, T_{ci+2}, T_{ci+3}, T_{ci+4}, T_{ci+5}, T_{ci+6}$, and T_{wi+7} . As is obvious in Fig. 1 along any base level there are cooling (heating) durations L_c (L_w) of various lengths. In the classical run theory of the statistics these durations are named as the positive and negative run lengths (Cramer and Leadbetter 1967). For a given temperature time series the number of cooling periods is either equal to the number of heating periods or the difference between these two numbers is equal to one (Şen 1977). The cooling (heating) durations show the continuation of uninterrupted cooling (heating) degree-days and they have the according time units, such as hours, days, months, or years.

(c) For any given temperature time series and base level there is a single maximum cooling (heating) duration, which is also referred to as the critical duration (Şen 1976). Critical duration is defined as the longest duration of uninterrupted cooling or heating successions in a given record length. It is logical to obtain the critical cooling (heating) duration change with the base level as a monotonic function. The cooling period decreases with the increasing level; the opposite is valid for heating periods.

(d) For any cooling period the summation of the cooling amounts is referred to as the cooling degree-day, D_c , and if the length of cooling duration is m then, in general,

$$D_c = \sum_{j=1}^m T_{c(i+j)}. \quad (3)$$

In this equation, i indicates the initial time of the cooling period. For time intervals less than one year i is season dependent. Similarly, the heating degree-day is defined as follows:

$$D_h = \sum_{j=1}^m T_{w(i+j)}. \quad (4)$$

These definitions can be calculated empirically from a given temperature time series, provided the various base levels are adopted. The variations in these quantities can be evaluated statistically by calculating their low-order moments, such as the arithmetic mean value, standard deviation, etc.

3. Theoretical approach

Theoretical considerations of cooling and heating properties can be achieved through the probability theory application. In general, the probability of a cooling period L_c with a length at least equal to m can be written implicitly as

$$P(L_c \geq m) = P(T_i \geq T_b, T_{i+1} \geq T_b, \dots, T_{i+j} \geq T_b, \dots, T_{i+m-1} \geq T_b, T_{i+m} \geq T_b). \quad (5)$$

A simpler version of this equation emerges when the successive occurrences of the temperature records are independent from each other:

$$P(L_c \geq m) = \prod_{j=1}^m P(T_{i+j} \geq T_b). \quad (6)$$

On the other hand, if the temperature occurrences are also identically distributed, then the simplest case results in the power form as

$$P(L_c \geq m) = [P(T_i \geq T_b)]^m. \quad (7)$$

If the probability distribution function of the temperature is known and indicated by $f(T)$, then

$$P(T_i \geq T_b) = \int_{T_b}^{\infty} f(T) dT = p. \quad (8)$$

The substitution of Eq. (8) into Eq. (7) leads to

$$P(L_c \geq m) = p^m. \quad (9)$$

On the other hand, Feller (1957) has proven the probability of a positive run length to be equal to an integer value m , in general, as

$$P(L_c = m) = P(L_c \geq m) - P(L_c \geq m + 1). \quad (10)$$

Thus, the probability of a cooling duration of length m can be obtained by substituting Eq. (9) into Eq. (10), which leads to

$$P(L_c = m) = qp^{m-1}, \quad (11)$$

where $q = 1 - p$. This last expression is valid for identically and independently distributed temperature variables only. However, the derivation of similar formulations is rather cumbersome for dependent variables. Şen (1977) has shown that if the successive data values are related to each other according to the first-order Markov process, then the probability statement similar to Eq. (7) becomes

$$P(L_c \geq m) = [P(T_i \geq T_b/T_{i-1} \geq T_b)]^m, \quad (12)$$

where the conditional probability within the square brackets represents the correlation coefficient. For the lag-one Markov process Şen (1976) has presented the following relationships that also hold for temperature distributions that comply with the Gaussian probability distribution function:

$$P(T_i \geq T_b/T_{i-1} \geq T_b) = r, \quad (13)$$

$$P(T_i \geq T_b/T_{i-1} \leq T_b) = \frac{p}{q}(1 - r), \quad (14)$$

$$P(T_i \leq T_b/T_{i-1} \leq T_b) = 1 - \frac{p}{q}(1 - r), \quad (15)$$

and

$$P(T_i \leq T_b/T_{i-1} \geq T_b) = 1 - r. \quad (16)$$

TABLE 1. The r values for a given base temperature percentage and ρ .

ρ	Percentage base temperatures				
	0.3	0.4	0.5	0.6	0.7
0.0	0.700	0.600	0.500	0.400	0.300
0.1	0.718	0.625	0.532	0.437	0.340
0.2	0.734	0.650	0.564	0.475	0.382
0.3	0.754	0.676	0.597	0.519	0.426
0.4	0.774	0.703	0.631	0.555	0.472
0.5	0.795	0.732	0.666	0.598	0.520
0.6	0.818	0.762	0.704	0.643	0.574
0.7	0.844	0.796	0.747	0.694	0.632
0.8	0.873	0.835	0.795	0.752	0.701
0.9	0.911	0.884	0.856	0.826	0.788

By definition $P(T_i < T_b/T_{i-1} < T_b)$ has been given in an integral form by Cramer and Leadbetter (1967) as

$$P(T_i \leq T_b/T_{i-1} \leq T_b) = q + \frac{1}{2\pi q} \int_0^{\rho} \exp[-x_0^2/2(1+z)](1-z^2)^{1/2} dz, \quad (17)$$

where ρ is the first-order correlation coefficient of the Markov process, which is given in its simplest form as

$$T_i - \bar{T} = \rho(T_{i-1} - \bar{T}) + \sigma\sqrt{1-\rho^2}\varepsilon_i, \quad (18)$$

where \bar{T} and σ are, respectively, the arithmetic mean and standard deviation values of the temperature data. Finally, ε_i s are the independent standard normal variates with mean zero and variance equal to one. The numerical solutions of Eq. (17) are given in Table 1 for various base temperature percentages, which are defined as the probability of T_i being less than T_b ; that is, it is equivalent to q . In light of the aforementioned calculations similar to the independent process case, the cooling duration probabilities can be written for dependent cases succinctly as

$$P(L_c \geq m) = r^{m-1}, \quad (19)$$

and, furthermore, by using Eq. (10), one can obtain from Eq. (19) that

$$P(L_c = m) = (1 - r)r^{m-1}. \quad (20)$$

The numerical solution of Eq. (20) has been obtained through computer software by considering some of the parameter sets in Table 1. The results are shown in Figs. 2–6. These graphs are the analytical solutions of the degree-day duration probabilities of any desired period, m , and for different correlation coefficients and base temperature percentages. Visual inspection and mutual comparison of these figures yield the following significant conclusions.

- 1) All the theoretical probability distribution functions of the cooling (heating) duration are negatively exponential. For longer periods the increase in serial correlation gives rise to higher probabilities.

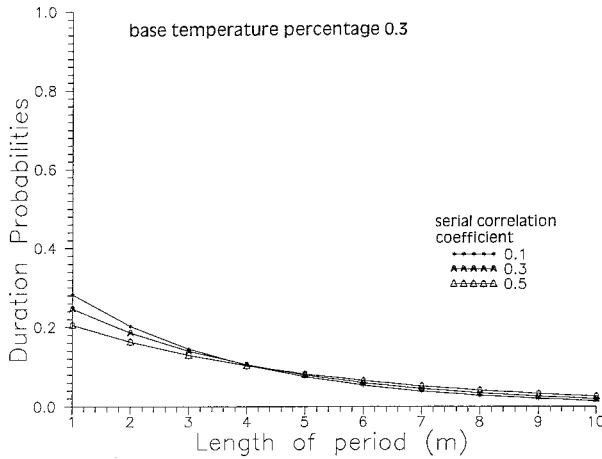


FIG. 2. Degree-day duration probability curves for base temperature percentages 0.3.

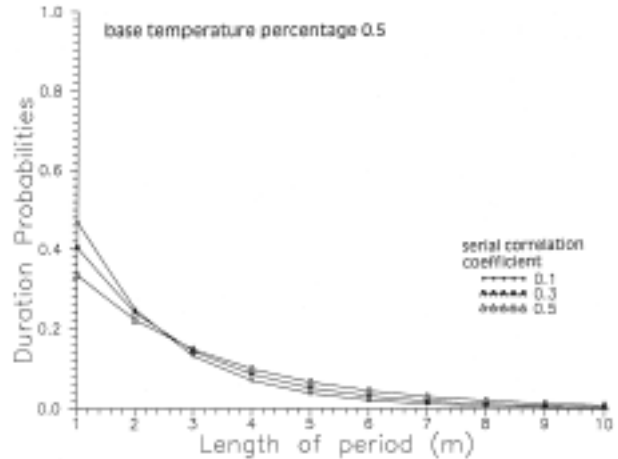


FIG. 4. Degree-day duration probability curves for base temperature percentages 0.5.

- 2) For the same base temperature an increase in the serial correlation coefficient gives rise to an increase in the cooling (heating) duration probability.
- 3) For long temperature time series, the probabilities of cooling (heating) durations merge asymptotically toward very small probability values. It is possible to conclude that dependence is not significant in duration calculations for large periods. This is tantamount to saying that independent process formulations are equally valid for dependent process cooling (heating) duration probability calculations for long periods. Furthermore Eqs. (11) and (20) converge for large m values.
- 4) For a given base temperature percentage and short periods, as the dependence increases in a temperature sequence, the probability of cooling (heating) duration increases.
- 5) Figures 2–6 show an increased probability for increased base temperature percentage. As the base

- 6) Contrary to item 3 above, the same probabilities increase for short periods.

An inspection of cooling degree-day definition in Eq. (3) indicates that D_c is a random variable consisting of two other random variables, namely, an integer random variable m corresponding to the duration of cooling period and a truncated but continuous random variable $T_{c(i+j)}$. Hence, it is logical to expect a joint probability distribution function between these two random variables as $P(m, D_c) = P(D_c/m)P(m)$, where $P(D_c/m)$ is the conditional probability of cooling degree day given its length, m . In fact, $P(m) = P(L_c = m)$, as explicitly given in Eq. (20). To find the marginal probability distribution function $P(D_c)$ of the cooling degree-day it is necessary and sufficient to sum all the possible cooling day durations theoretically as

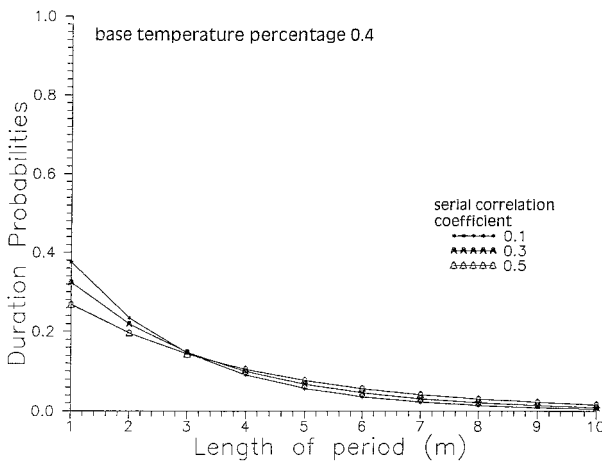


FIG. 3. Degree-day duration probability curves for base temperature percentages 0.4.

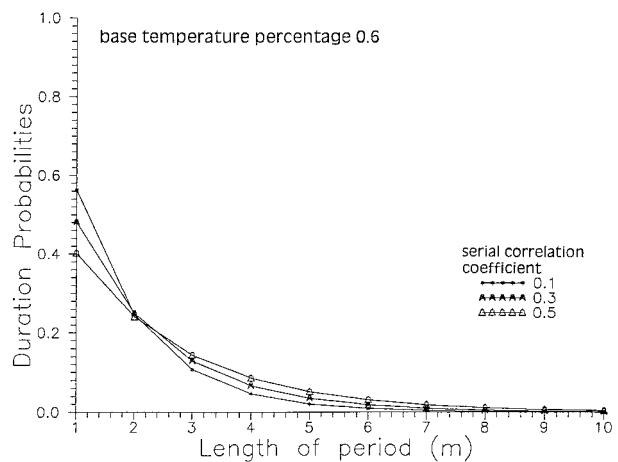


FIG. 5. Degree-day duration probability curves for base temperature percentages 0.6.

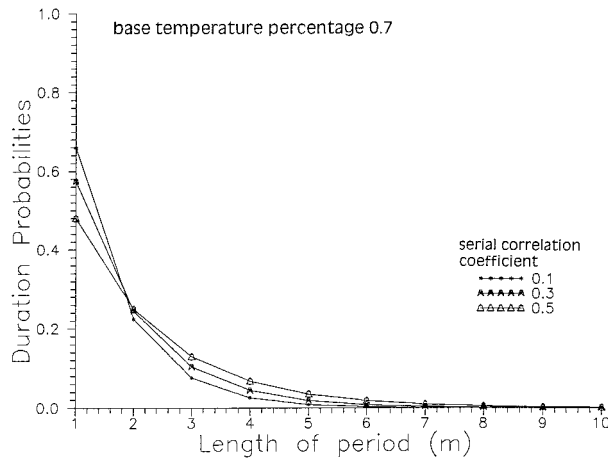


FIG. 6. Degree-day duration probability curves for base temperature percentages 0.7.

$$P(D_c) = \sum_{m=0}^{\infty} P(D_c/m)P(m). \quad (21)$$

Hence, the cooling degree-day expectation becomes from Eq. (21) as

$$E(D_c) = E(L_c)E(T_c). \quad (22)$$

4. Applications

The methodology explained in the previous section is applied for five stations in various parts of Turkey, as shown in Fig. 7. Daily temperature data are taken from each site, selected as representatives of different climatic regions in Turkey. On this basis, Izmir represents the Aegean region; Adana is representative of the Mediterranean climate; continental climatic conditions are characteristic in the capital of Turkey, Ankara; the Black Sea features are reflected in the coastal city of Samsun along the Black Sea coast; and Kars in eastern Turkey is frequently under the influence of Siberian air

masses, especially during the winter season. The significant statistical properties of each station are represented in Table 2. At first glance, it is possible to see from this table that the average and median values are almost equal to each other within 5% relative error limit, and, consequently, the temperature distribution functions are approximately Gaussian. Furthermore, this point is confirmed by the very small skewness coefficients in the same table. The average and median values together with the standard deviations have comparatively larger values at cities along the sea coasts, but in continental locations, such as Ankara and Kars, they are smaller.

From long records of measurements and experience, cooling and heating days in Turkey have been found to be represented by base levels of 24° and 15°C, respectively. These bases are assumed equally valid for each subclimatic region within the country in this study, although it is possible to adopt any other desired truncation level. Figures 8 and 9 indicate annual average heating and cooling degree-day, and relative frequency distributions for each one of the cities considered, respectively. They are all in the form of a gamma distribution. This distribution subsumes negative exponential distribution as one of its special cases. The goodness of fit between the experimental and theoretical relative frequency distributions is achieved by the commonly used χ^2 test. This test is based upon how good of a fit there is between the frequency of observations occurrence in a sample and the expected frequencies from the hypothesized theoretical distribution. In general, the goodness-of-fit test between observed, O_i , and expected, E_i , frequencies with k frequency groups is defined as

$$\aleph^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}, \quad (23)$$

where \aleph^2 is a value of the random variable whose sampling distribution is approximated very closely by the χ^2 distribution. In the case of closely observed frequencies to the corresponding expected frequencies, the

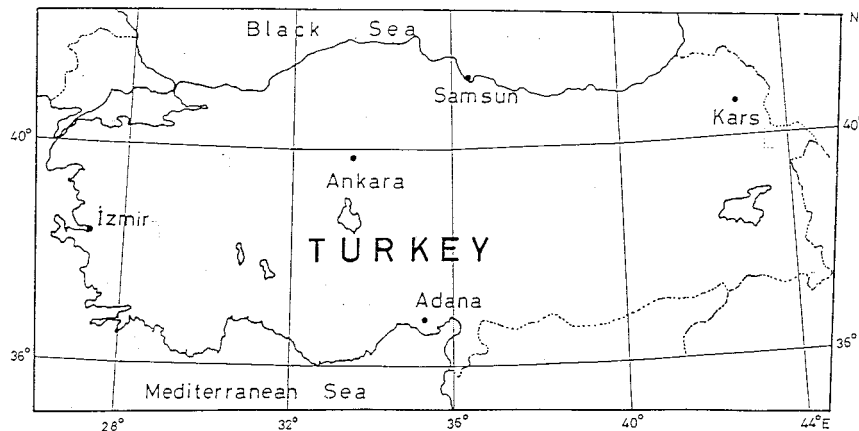


FIG. 7. Location map.

TABLE 2. Statistical properties of the stations.

Variable	İzmir	Adana	Ankara	Samsun	Kars
Sample size	3655.0	3655.0	3655.0	3655.0	3655.0
Average (°C)	17.3	18.5	12.3	14.7	8.9
Median (°C)	17.2	19.0	12.8	15.1	9.7
Mode (°C)	27.2	28.1	21.6	21.1	0.0
Variance (°C ²)	56.8	53.5	73.7	41.2	50.6
Standard deviation (°C)	7.5	7.3	8.6	6.4	7.1
Standard error (°C)	0.1	0.1	0.1	0.1	0.1
Minimum (°C)	-2.7	-0.9	-7.4	-1.8	-22.8
Maximum (°C)	33.2	32.8	29.9	29.4	25.2
Range (°C)	35.9	33.7	37.3	31.2	48.0
Lower quartile (°C)	11.1	12.1	4.7	9.7	3.0
Upper quartile (°C)	24.1	25.5	19.7	20.3	14.7
Interquartile range (°C)	13.0	13.4	15.0	10.6	11.7
Skewness (°C ³)	-0.1	-0.2	-0.1	-0.2	-0.3
Standardized skewness	-1.7	-3.8	-2.9	-6.0	-6.5
Kurtosis (°C ⁴)	-1.0	-1.2	-1.1	-0.9	-0.7
Standardized kurtosis	-13.0	-14.6	-13.5	-11.4	-8.8

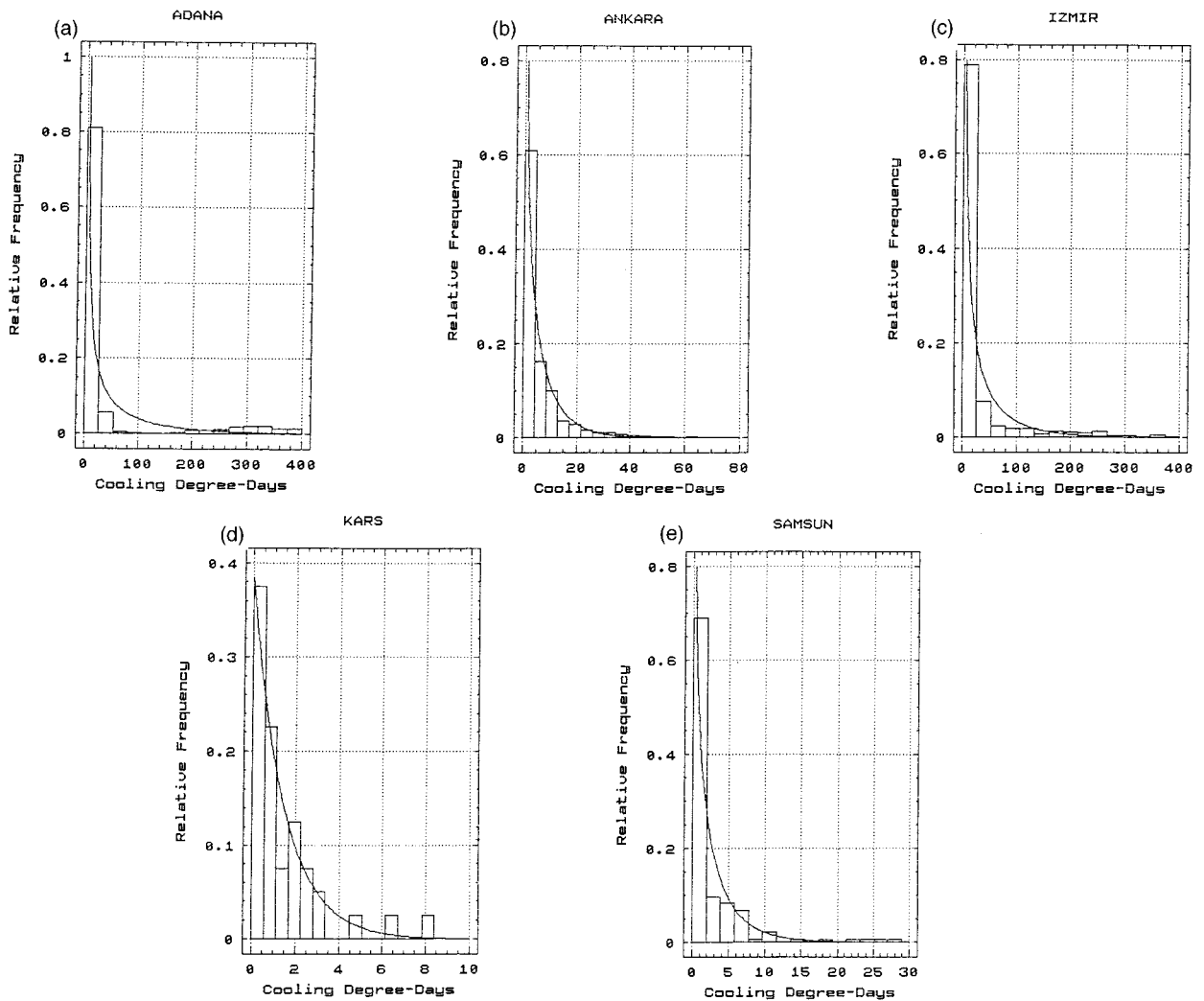


FIG. 8. Annual average cooling degree-days (°C).

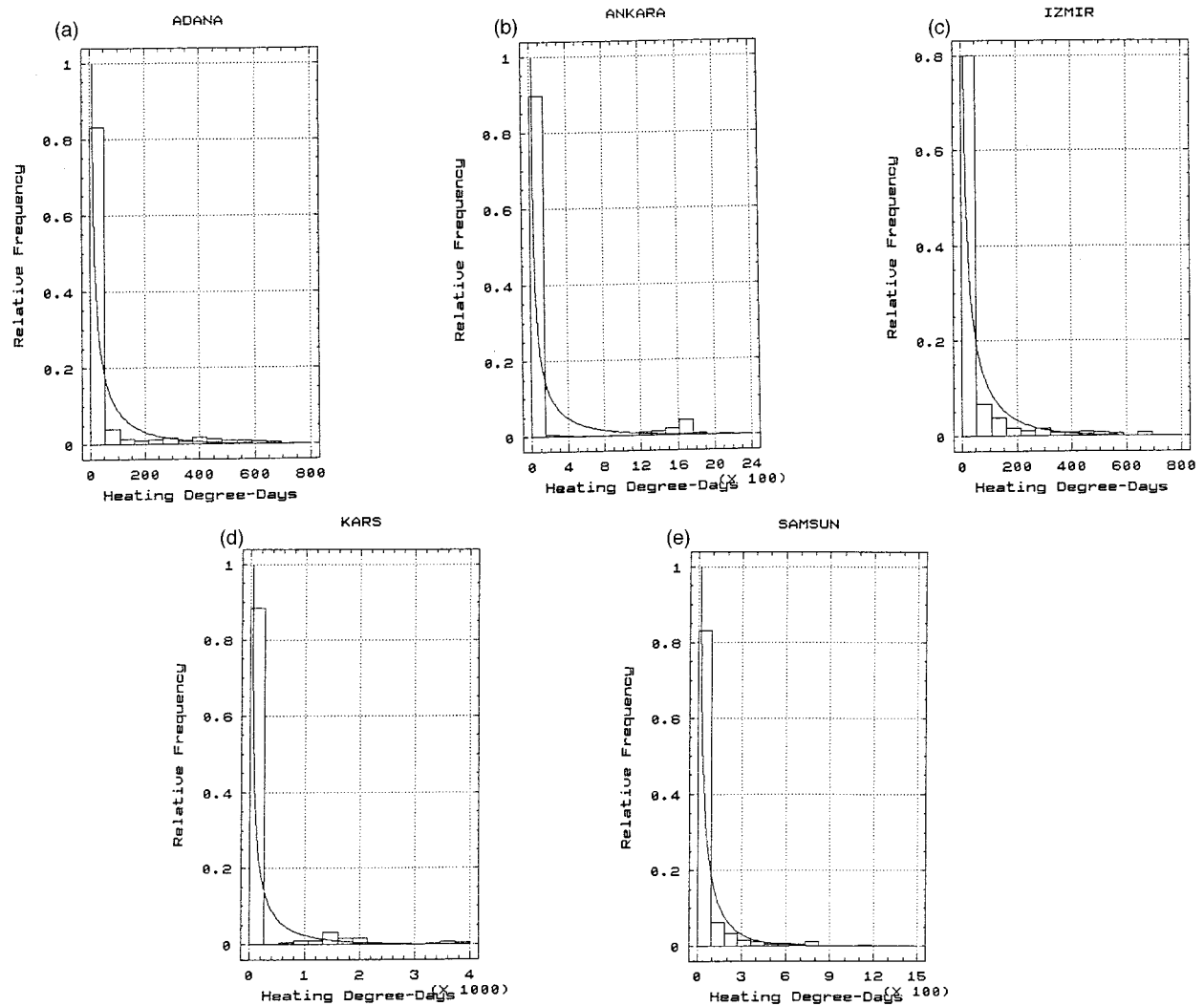


FIG. 9. Annual average heating degree-days (°C).

χ^2 value will be small, indicating a good fit. The χ^2 test on one side at a 5% significance level is used in this paper in order to examine the combinations or contingency of the observed and expected frequencies and finally to make the decision about the type of best distribution for the data at hand. First of all, for given data and relevant degrees of freedom, a critical χ^2 value is

found at 5% significance level from χ^2 distribution tables in any statistics textbook. It is observed that χ^2 values are all less than the critical level. A gamma distribution implies that moderately lower degree-days occur more frequently than higher degree-days in Turkey, irrespective of heating or cooling, as compared to many parts of the world. However, relative positions of the stations show that from the west toward the east cooling (heating) degree-day frequency decreases (increases) significantly reaching a minimum (maximum) value at Kars. Table 3 presents the absolute heating and cooling degree-day frequencies for the smallest degree-day classes in addition to the relative difference percentages between the heating and cooling degree days at each station. Relative difference percentage is defined as the absolute difference between the smallest class frequencies of heating and cooling degree-days divided by the greatest of these two frequencies multiplied by 100.

TABLE 3. Degree-day (°C) properties.

Station	Heating		Cooling		Relative difference (%)
	Smallest class	Frequency	Smallest class	Frequency	
İzmir	0-23	410	0-50	540	24
Adana	0-23	410	0-50	475	14
Ankara	0-4	410	0-1.5	612	48
Samsun	0-2	320	0-1	410	70
Kars	0-2	30	0-0.22	355	91

TABLE 4. Annual average degree-days.

Station	Heating	Cooling	Relative difference (%)
İzmir	648	372	43
Adana	648	400	38
Ankara	1655	62.5	96
Samsun	1400	30	98
Kars	1400	10	99

It is clear that the differences between the heating and cooling days increase when the continental type of climatic effects become more pronounced, as in the Kars case. Smaller relative differences represent more temperate climates of the region. High relative difference percentage locations need more heating than cooling. Annual average heating and cooling degree-day maximum values were extracted from Figs. 8 and 9, leading

to the values in Table 4. These values have minimum probabilities of occurrences. As one moves toward continental and northern subtropical climate regions, the heating degree-day increases tremendously, whereas the cooling degree-days decrease.

Figures 10 and 11 indicate the relative frequency distributions of heating and cooling durations, both in the form of gamma distributions. These are well in accordance with the theoretical probability density functions in Figs. 2–6. Longer periods, in general, show more persistent durations, which is tantamount to saying that successive interrupted cooling or heating day durations are longer than usual. One interesting practical problem is to seek the heating or cooling spells that exhibit the most severe conditions in terms of energy requirements. More specifically, the question is whether the association of long durations with weak or severe degree-days with short durations are more demanding of energy. A good definitive answer to these questions can be ob-

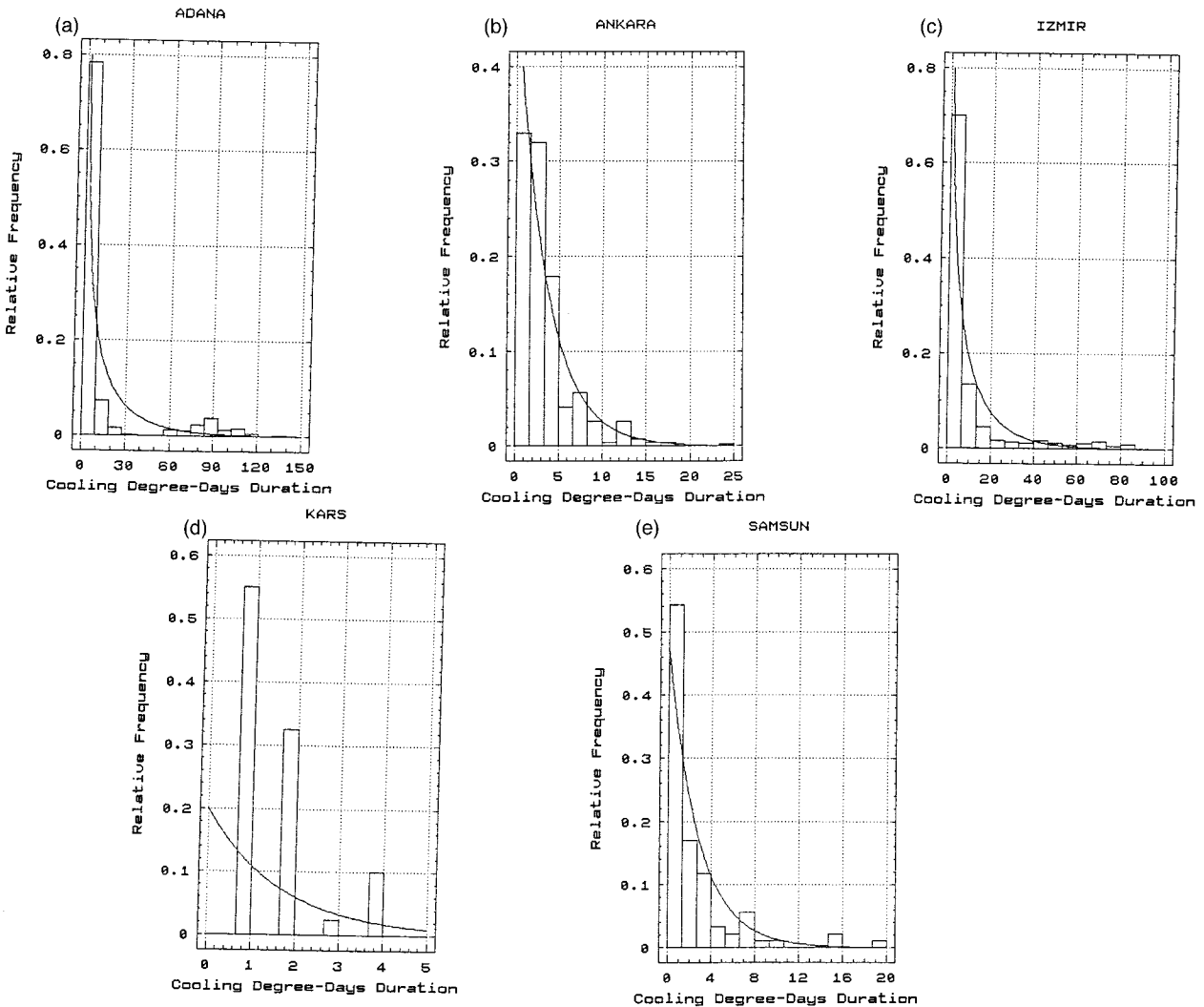


FIG. 10. Cooling degree-days duration.

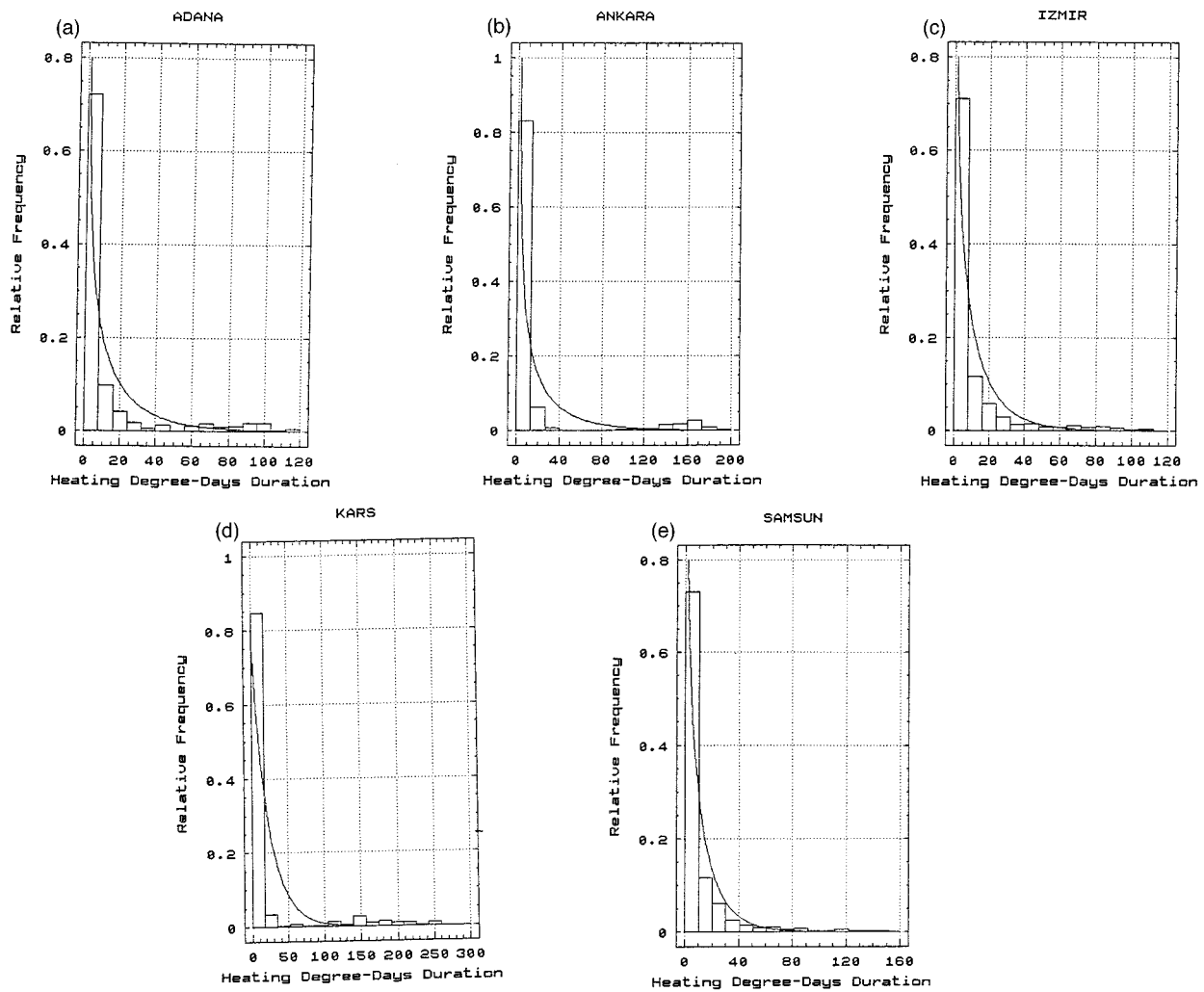


FIG. 11. Heating degree-days duration.

tained objectively by multiplication of average heating or cooling days with their respective average durations, as expressed in Eq. (22).

5. Conclusions

Definitions of degree-day concepts for cooling and heating periods are presented in terms of simple terminology followed by a detailed analytical derivation of these properties for dependent first-order Markov processes on the basis of probability theory. General formulations for various degree-day definitions are derived by considering the dependence within the temperature time series. Necessary tables are presented, and subsequently numerical solutions of formulations for different base temperatures and serial correlation coefficients are presented in graphic forms.

All of these graphs indicate that the cooling and heating durations are distributed according to a gamma (neg-

ative exponential) distribution. Also, any increase in the serial correlation of the data gives rise to increases in these durations especially for small periods.

Finally, five representative temperature records are adopted from different climate regions of Turkey and their degree-day performances are documented by tables and figures. It is observed that in regions of continental climate within Turkey, differences between the heating and cooling degree-days increase.

REFERENCES

Cramer, H., and M. R. Leadbetter, 1967: *Stationary and Related Stochastic Processes*. John Wiley, 384 pp.
 Feller, W., 1957: *An Introduction to the Probability Theory and Its Application*. John Wiley, 509 pp.
 Guttman, N. B., 1983: Variability of population-weighted seasonal heating degree days. *J. Climate Appl. Meteor.*, **22**, 495-501.
 —, and R. L. Lehman, 1992: Estimation of daily degree hours. *J. Appl. Meteor.*, **31**, 797-810.

- Lehman, R. L., 1987: Probability distribution of monthly degree day variables at U.S. stations. Part 1: Estimating the mean value and variance from temperature data. *J. Appl. Probab.*, **26**, 329–340.
- Quayle, R. G., and H. F. Diaz, 1980: Heating degree day data applied to residential heating energy consumption. *J. Appl. Meteor.*, **19**, 241–246.
- Şen, Z., 1976: Wet and dry periods of annual flow series. *J. Hydraul. Div., Amer. Soc. Civ. Eng.*, **102**, 1503–1514.
- , 1977: Run sums of annual flow series. *J. Hydrol.*, **35**, 311–324.
- Thom, H. S. C., 1952: Seasonal degree day statistics for the United States. *Mon. Wea. Rev.*, **80**, 143–149.
- , 1954: The rational relationship between heating degree days and temperature. *Mon. Wea. Rev.*, **82**, 1–6.
- , 1966: Normal degree days above any base by the universal truncation coefficient. *Mon. Wea. Rev.*, **94**, 461–465.