Available online at www.sciencedirect.com
SCIENCE DIRECT®
Energy Conversion and Management 47 (2006) 1843–1850
ENERGY CONVERSION & MANAGEMENT
www.elsevier.com/locate/enconman

A stochastic method to generate bin weather data in Nanjing, China

Zhou Jin, Wu Yezheng *, Yan Gang

Department of Refrigeration and Cryogenic Engineering, School of Energy and Power Engineering, Xi'an Jiao Tong University, Xi'an 710049, PR China

Received 4 April 2005; accepted 2 October 2005
Available online 1 December 2005

Abstract

The bin method, as one of the well known and simple steady state methods used to predict heating and cooling energy consumption of buildings, requires reliable and detailed bin data. Since the long term hourly temperature records are not available in China, there is a lack of bin weather data for study and use. In order to keep the bin method practical in China, a stochastic model using only the daily maximum and minimum temperatures to generate bin weather data was established and tested by applying one year of measured hourly ambient temperature data in Nanjing, China. By comparison with the measured values, the bin weather data generated by the model shows adequate accuracy. This stochastic model can be used to estimate the bin weather data in areas, especially in China, where the long term hourly temperature records are missing or not available.

Keywords: Energy analysis; Stochastic method; Bin data; China

1. Introduction

In the sense of minimizing the life cycle cost of a building, energy analysis plays an important role in developing an optimum and cost effective design of a heating or an air conditioning system for a building. Several models are available for estimating energy use in buildings. These models range from simple steady state models to comprehensive dynamic simulation procedures.

Today, several computer programs, in which the influence of many parameters that are mainly functions of time are taken into consideration, are available for simulating both buildings and systems and performing hour by hour energy calculations using hourly weather data. DOE-2, BLAST and TRNSYS are such

* Corresponding author. Tel.: +86 29 8266 8738; fax: +86 29 8266 8725.
E-mail address: yzwu@mail.xjtu.edu.cn (W. Yezheng).

0196-8904/$ - see front matter
doi:10.1016/j.enconman.2005.10.006

Nomenclature

D	number of days
F	frequency of normalized hourly ambient temperature
MAPE	mean absolute percentage error (%)
n	number of subintervals into which the interval [0, 1] was equally divided
N	number of normalized temperatures that fall in subinterval
p	probability density
t	hourly ambient temperature (°C)
$\hat{t}$	normalized hourly ambient temperature (dimensionless)
W	weighting factor

Subscripts

c	calculated value
m	measured value
max	daily maximum
min	daily minimum

programs that have gained widespread acceptance as reliable estimation tools. Unfortunately, along with the increased sophistication of these models, they have also become very complex and tedious to use [1].

The steady state methods, which are also called single measure methods, require less data and provide adequate results for simple systems and applications. These methods are appropriate if the utilization of the building can be considered constant. Among these methods are the degree day and bin data methods.

The degree-day methods are the best known and the simplest methods among the steady state models. Traditionally, the degree-day method is based on the assumption that on a long term average, the solar and internal gains will offset the heat loss when the mean daily outdoor temperature is 18.3 °C and that the energy consumption will be proportional to the difference between 18.3 °C and the mean daily temperature. The degree-day method can estimate energy consumption very accurately if the building use and the efficiency of the HVAC equipment are sufficiently constant. However, for many applications, at least one of the above parameters varies with time. For instance, the efficiency of a heat pump system and HVAC equipment may be affected directly or indirectly by outdoor temperature. In such cases, the bin method can yield good results for the annual energy consumption if different temperature intervals and time periods are evaluated separately. In the bin method, the energy consumption is calculated for several values of the outdoor temperature and multiplied by the number of hours in the temperature interval (bin) centered around that temperature. Bin data is defined as the number of hours that the ambient temperature was in each of a set of equally sized intervals of ambient temperature.

In the United States, the necessary bin weather data are available in the literature [2,3]. Some researchers [4–8] have developed bin weather data for other regions of the world. However, there is a lack of information in the ASHRAE handbooks concerning the bin weather data required to perform energy calculations in buildings in China. The practice of analysis of weather data for the design of HVAC systems and energy consumption predictions in China is quite new. For a long time, only the daily value of meteorological elements, such as daily maximum, minimum and average temperature, was recorded and available in most meteorological observations in China, but what was needed to obtain the bin weather data, such as temperature bin data, were the long term hourly values of air temperature. The study of bin weather data is very limited in China. Only a few attempts [9,10] in which bin weather data for several cities was given have been found in China. Obviously, this cannot meet the need for actual use and research. So, there is an urgent need for developing bin weather data in China. The objective of this paper, therefore, is to study the hourly measured air temperature distribution and then to establish a model to generate bin weather data for the long term daily temperature data.

2. Data used

In this paper, to study the hourly ambient temperature variation and to establish and evaluate the model, a one year long hourly ambient temperature record for Nanjing in 2002 was used in the study. These data are taken from the Climatological Center of Lukou Airport in Nanjing, which is located in the southeast of China (latitude 32.0°N, longitude 118.8°E, altitude 9 m).

In addition, in order to create the bin weather data for Nanjing, typical weather year data was needed. Based on the long term meteorological data from 1961 to 1989 obtained from the China Meteorological Administration, the typical weather year data for most cities in China has been studied in our former research [11] by means of the TMY (Typical Meteorological Year) method. The typical weather year for Nanjing is shown in Table 1. As only daily values of the meteorological elements were recorded and available in China, the data contained in the typical weather year data was also only daily values. In this study, the daily maximum and minimum ambient temperature in the typical weather year data for Nanjing was used.

3. Stochastic model to generate bin data

Traditionally, the generation of bin weather data needs long term hourly ambient temperature records. However, in the generation, the time information, namely the exact time that such a temperature occurred in a day, was omitted, and only the numerical value of the temperature was used. So, the value of each hourly ambient temperature can be treated as the independent random variable, and its distribution within the daily temperature range can be analyzed by means of probability theory.

3.1. Probability distribution of normalized hourly ambient temperature

Since the daily maximum and minimum temperatures and temperature range varied day by day, the concept of normalized hourly ambient temperature should be introduced to transform the hourly temperatures in each day into a uniform scale. The new variable, normalized hourly ambient temperature is defined by

$$\hat{t} = \frac{t - t_{\min}}{t_{\max} - t_{\min}} \tag{1}$$

where $\hat{t}$ may be termed the normalized hourly ambient temperature, $t_{\max}$ and $t_{\min}$ are the daily maximum and minimum temperatures, respectively, t is the hourly ambient temperature.

Obviously, the normalized hourly ambient temperature $\hat{t}$ is a random variable that lies in the interval $[0, 1]$. To analyze its distribution, the interval $[0, 1]$ can be divided equally into several subintervals, and by means of the histogram method [12], the probability distribution density of $\hat{t}$ in each subinterval can be calculated by

$$p_i = \frac{F_i}{1/n} \tag{2}$$

Table 1
The typical weather year for Nanjing

Month	Year
January	1982
February	1987
March	1983
April	1987
May	1975
June	1964
July	1984
August	1968
September	1986
October	1984
November	1989
December	1982

where p_i is the probability density of the ith subinterval, n is the number of subintervals and F_i is the frequency that $\hat{t}$ lies in the ith subinterval. For a whole day period, F_i can be calculated by

$$F_i = \begin{cases} (N_i - D)\Big/\left(\sum_{i=1}^{n} N_i - 2\cdot D\right), & i = 1, n \\ N_i\Big/\left(\sum_{i=1}^{n} N_i - 2\cdot D\right), & i \neq 1, n \end{cases} \tag{3}$$

where N_i is the number of normalized temperatures that fall in ith subinterval, D is the number of days that the data was calculated. In addition, other periods in a day, such as the period 08:00–20:00, are often required in the generation of bin weather data. In this case, F_i can be obtained by

$$F_i = \begin{cases} N_{i+1}\Big/\left[\sum_{i=1}^{n} N_i - (N_n - N_{n-1}) - (N_1 - N_2)\right], & i = 1 \\ N_i\Big/\left[\sum_{i=1}^{n} N_i - (N_n - N_{n-1}) - (N_1 - N_2)\right], & i \neq 1, n \\ N_{i-1}\Big/\left[\sum_{i=1}^{n} N_i - (N_n - N_{n-1}) - (N_1 - N_2)\right], & i = n \end{cases} \tag{4}$$

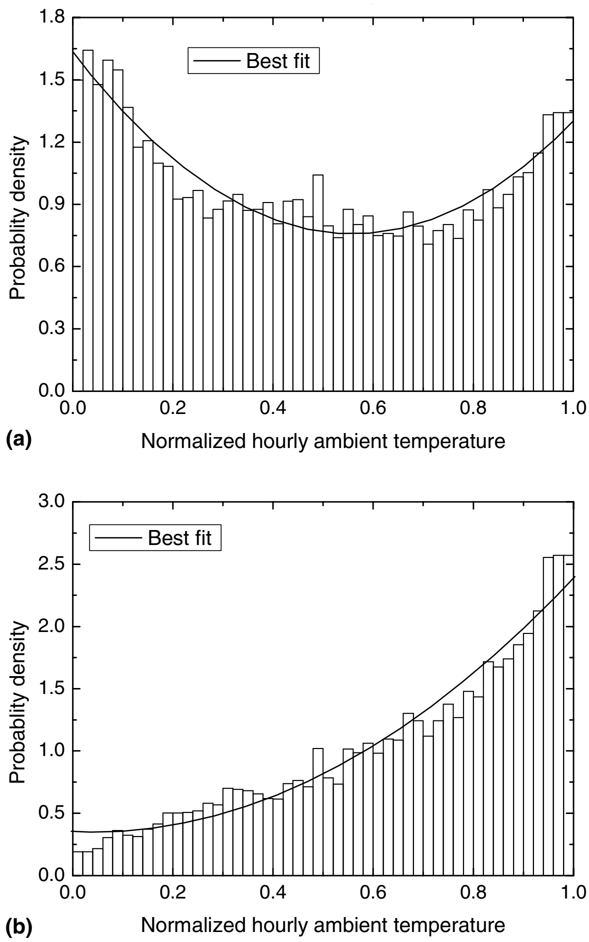

Fig. 1. (a) Whole day period. (b) 08:00–20:00 period. Fitting of probability density function of normalized hourly ambient temperature.

Based on the one year long hourly ambient temperature data in Nanjing, China, the probability density p_i was calculated for the whole day and the 08:00–20:00 period, where the interval $[0,1]$ was equally divided into 50 subintervals, namely n equals 50. The results are shown in Fig. 1.

According to the discrete probability density data in Fig. 1, the probability density function of $\hat{t}$ can be obtained by a fitting method. In this study, the quadratic polynomials

$$p = a\hat{t}^2 + b\hat{t} + c, \quad 0 \leqslant \hat{t} \leqslant 1 \tag{5}$$

were used to fit the probability density data, where a, b and c are coefficients. According to the property of the probability density function, the following equation should be satisfied

$$\int_0^1 (a\hat{t}^2 + b\hat{t} + c)\mathrm{d}\hat{t} = \frac{a}{3} + \frac{b}{2} + c = 1 \tag{6}$$

As shown in Fig. 1, the probability density curve obtained according to the probability density data points is also shown. The probability density functions that are fitted are described by

$$p = 2.7893\hat{t}^2 - 3.1228\hat{t} + 1.6316 \quad \text{for the whole day period} \tag{7}$$

$$p = 2.2173\hat{t}^2 - 0.1827\hat{t} + 0.3522 \quad \text{for the 08 : 00–20 : 00 period} \tag{8}$$

3.2. The generation of hourly ambient temperature

As stated in the beginning of this paper, the objective of this study is to generate the hourly ambient temperature needed for bin weather data generation in the case that only the daily maximum and minimum temperatures are known. To do this, we can use the obtained probability density function to generate the normalized hourly ambient temperature and then transform it to hourly temperature. This belongs to the problem of how to simulate a random variable with a prescribed probability density function and can be done on a computer by the method described in the literature [13]. For a given probability density function $f(\hat{t})$, if its distribution function $F(\hat{t})$ can be obtained and if u is a random variable with uniform distribution on $[0,1]$, then

$$\hat{t} = F^{-1}(u) \tag{9}$$

has the distribution $F(\hat{t})$. Thus, to simulate a random variable with a given distribution F, we need only set $\hat{t} = F^{-1}(u)$.

As stated above, the probability density function of the normalized ambient temperature was fitted using a one year long hourly temperature data. Based on the probability density function obtained, the random normalized hourly temperature can be generated. When the daily maximum and minimum temperature are known, the normalized hourly temperature can be transformed to an actual temperature by the following equation

$$t = \hat{t} \cdot (t_{\max} - t_{\min}) + t_{\min} \tag{10}$$

3.3. Obtaining the bin data

When the hourly temperature for a particular period of the day has been generated using the above method, the bin data can also be obtained. Because the normalized temperature generated using the model in this study is a random variable, the bin data obtained from each generation shows some difference, but it has much similarity. To obtain a stable result of bin data, the generation of the bin data can be performed enough times, and the bin data can be obtained by averaging the result of each generation. In this paper, 50 generations were averaged to generate the bin weather data.

3.4. Methods of model evaluation

The performance of the model was evaluated in terms of the following statistical error test:

$$\mathrm{MAPE} = \sum_i \left(\frac{|N_{i\cdot c} - N_{i\cdot m}|}{N_{i\cdot m}} \cdot W_i \right) \cdot 100 \tag{11}$$

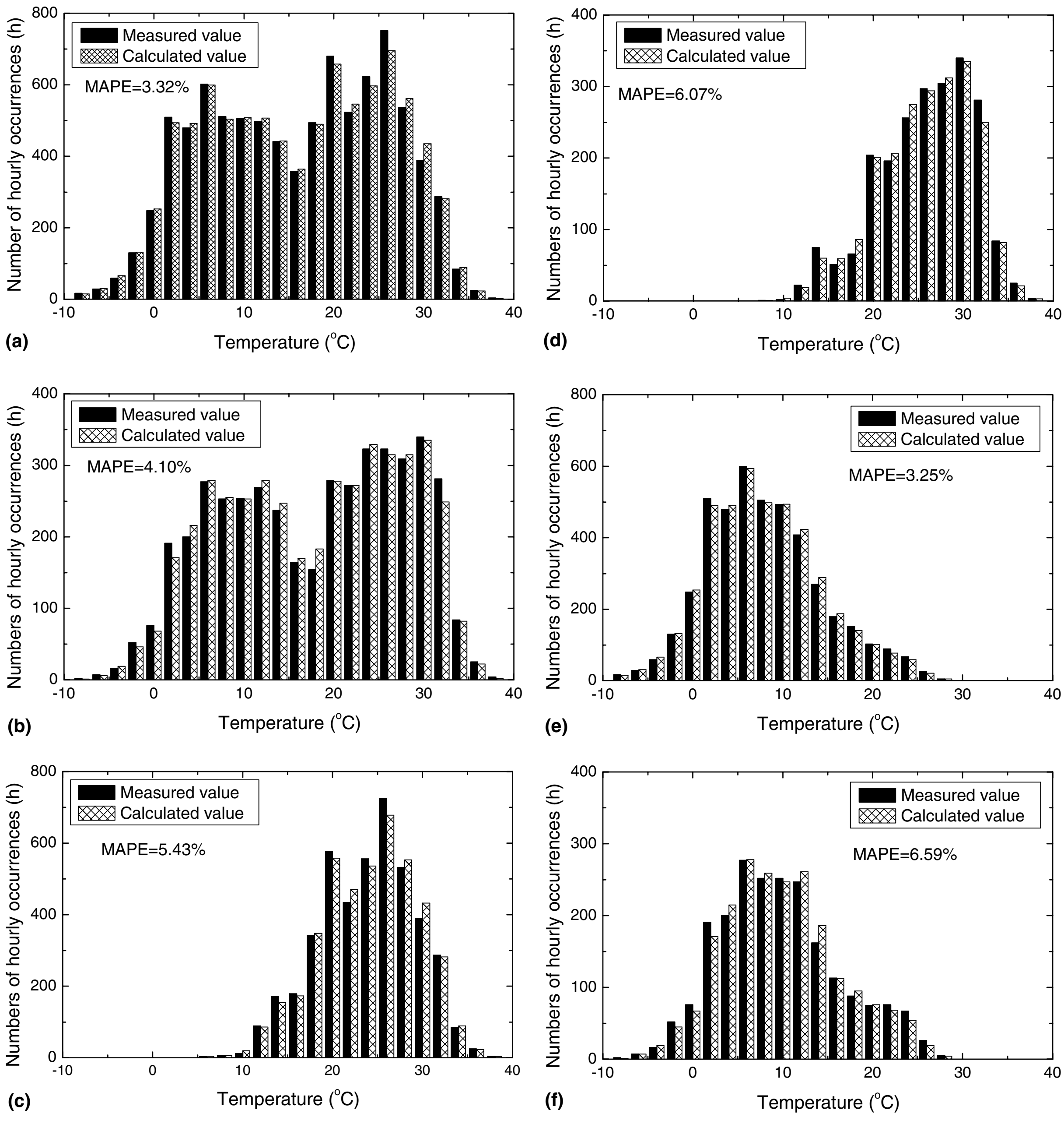

Fig. 2. (a) Total year, whole day 24 h period. (b) Total year, 08:00–20:00 period. (c) From May 1 to October 31, whole day 24 h period. (d) From May 1 to October 31, 08:00–20:00 period. (e) From November 1 to April 30, whole day 24 h period. (f) From November 1 to April 30, 08:00–20:00 period. The measured and calculated values of bin data at Nanjing in 2002.

where

$$W_i = \frac{N_{i\cdot\mathrm{m}}}{\sum_j N_{j\cdot\mathrm{m}}} \tag{12}$$

As the frequency distribution of the temperature is a bell shaped curve, it can be assumed that the central part of the data is the most meaningful and that the lowest and highest bins represent exceptional weather conditions. Since the bin method is used to calculate the energy consumption of buildings, it makes sense to focus on the central bins, which account for the major part of the energy consumption. In order to take this into account in the model evaluation, a weighting factor, W_i, is introduced in Eq. (11).

4. Results and discussion

4.1. Validation of the model

Based on the aforementioned model established in this study and the daily maximum and minimum temperature record at Nanjing in 2002, the bin weather data for two different seasons, including the period from May 1 to October 31 and the period from November 1 to April 30, was produced. For each season, two different periods of the day, including the whole day of 24 h and the 08:00–20:00 period, were considered. The bin weather data obtained are shown in Fig. 2, where the actual bin data and statistical error are also calculated and shown.

As shown in Fig. 2, the errors resulting from the model established in this paper lie between 3.25% and 6.59%. This result is considered acceptable. Although the comparison was only based on a one year long temperature record at Nanjing (and so the bin data in Fig. 2 cannot be used to calculate energy consumption of a

Table 2
Bin data in Nanjing

Temperature bin (°C)	Period					
	From May 1 to Oct 31		From Nov 1 to Apr 30		Total year	
	Whole day 24 h	08:00–20:00	Whole day 24 h	08:00–20:00	Whole day 24 h	08:00–20:00
−11/−9	0	0	3	0	3	0
−9/−7	0	0	12	1	12	1
−7/−5	0	0	20	2	20	2
−5/−3	0	0	113	13	113	13
−3/−1	0	0	214	38	214	38
−1/1	0	0	324	89	324	89
1/3	0	0	421	160	421	160
3/5	0	0	510	200	510	200
5/7	15	1	598	285	613	286
7/9	29	3	517	308	546	311
9/11	38	6	357	208	395	214
11/13	78	16	329	223	407	239
13/15	176	35	345	221	521	256
15/17	318	82	213	139	531	221
17/19	362	131	165	117	527	248
19/21	477	222	74	55	551	277
21/23	522	248	55	47	577	295
23/25	630	300	42	38	672	338
25/27	579	272	19	16	598	288
27/29	478	287	8	8	486	295
29/31	345	268	3	2	348	270
31/33	237	216	2	2	239	218
33/35	117	106	0	0	117	106
35/37	15	15	0	0	15	15

building), it indicates that the model in this study is feasible and can be used to generate bin data if there are long term temperature records available.

4.2. Bin data in Nanjing

The constitution of the typical weather year in most cities of China has been studied by the TMY method in our former research [11], in which only the daily maximum, minimum and average ambient temperatures was included. Based on the model established in this paper and the typical weather year data for Nanjing in Table 1, the bin weather data for two seasons and two periods of the day was obtained, and the result is shown in Table 2. As shown in Table 2, the temperature bin was set to be 2 °C in the generation of the bin data. In addition, the bin data for the night 20:00–08:00 period can be obtained by subtracting the data in each bin for the 08:00–20:00 period from that for the whole day if it is required.

5. Conclusion

Bin weather data is essential to the application of the bin method, which can be efficiently used to estimate the heating and cooling energy requirements of buildings. A stochastic model, in which only the daily maximum and minimum ambient temperatures were used, to generate bin weather data was established in this paper. Using this stochastic model, the bin data for Nanjing were generated. The statistical estimator MAPE was used to assess this model. For the different operating seasons and periods of occupancy considered in this paper, the MAPE ranges from 3.25% to 6.59%. This model can be used to estimate the bin weather data in areas where the hourly temperature data is missing or not available.

References

[1] Degelmen LO. Bin weather data for simplified energy calculations and variable degree-day information. ASHRAE Trans 1985;91(1A):3–14.
[2] ASHRAE. Bin and degree hour weather data for simplified energy calculations; 1995.
[3] USAF. Engineering weather data. Department of the Air Force Manual AFM 88-29. Washington, DC: US Government Printing Office 1978.
[4] Bulut H, Buyukalaca O, Yilmaz T. Bin weather data for Turkey. Appl Energ 2001;70(2):135–55.
[5] Orhan B, Husamettin B. Detailed weather data for the provinces covered by the Southeastern Anatolia Project GAP of Turkey. Appl Energ 2004;77(2):187–204.
[6] Papakostas KT. Bin weather data of Athens, Greece. Renew Energ 1999;17(2):265–75.
[7] Papakostas KT, Sotiropoulos BA. Bin weather data of Thessaloniki, Greece. Renew Energ 1997;11(1):69–76.
[8] Said SAM, Habib MA, Iqbal MO. Database for building energy prediction in Saudi Arabia. Energ Convers Manage 2003;44(1):191–201.
[9] Tian S. Simplified energy estimation method for HVAC in China. Proc of the annual meeting of HVAC in southwest China. Kunming, China; 1993.
[10] Li L. Bin method-simplified energy estimation method for HVAC in China. MPhil Thesis, Chongqing, China: Chongqing Jianzhu University; 1993.
[11] Zhou J. The origination of bin weather data and new solar load model based on the bin method. PhD thesis. Xi'an, China: Xi'an Jiaotong University; 2005.
[12] Moore DS, McCabe GP. Introduction to the practice of statistics. New York: Freeman; 1993.
[13] Laurie Snell J. Introduction to probability. New York: Random House; 1987.