CHAPTER 6

WEATHER DERIVATIVES

SECTIONS

The impact of weather on many commercial businesses and recreational activities is significant and varies both geographically and seasonally. Many businesses, including agriculture, insurance, energy, and tourism, are either favorably or adversely affected by weather. For this reason, the financial markets have created an innovative new class of instruments called *weather derivatives*, through which risk exposure to weather (and temperature) may be transferred or reduced. Weather derivatives are contingent claims written on weather indices, which in turn, are variables whose values are constructed from weather data. Commonly referenced weather indices include, but are not restricted to, Daily Average Temperature (DAT), Cumulative Annual Temperature (CAT), Heating Degree Days (HDDs), Cooling Degree Days (CDDs), precipitation, snowfall, and wind. In contrast to other contingent claims, weather derivatives pricing poses some difficulties because pricing based on the traditional risk-neutral no-arbitrage arguments do not work; underlying weather indices at present are not securitized by liquid traded instruments. There also exist some difficulties in implementing statistical, equilibrium-based pricing techniques, because the observed weather indices are non-stationary, and "characterized by long-term variations and trends, potentially with cycles much longer than what the data records reveal."[1] In contrast, an actuarial present value pricing approach is rather simple and intuitively appealing, although it cannot capture such cycles and trends in the weather like statistical models. In this chapter, we explore various pricing models for weather derivatives, including use of statistical and stochastic models.

In §6.1, we provide a background of the weather derivatives market and how it originated. In §6.2, we discuss weather contracts, including CME futures contracts and options contracts on weather indices. In §6.3, we cover modeling temperature based on the work of Alaton et al. (2000). In §6.4, we discuss general parameter estimation of the model in §6.3. In §6.5, we review mean-reversion estimation, while in §6.6, we discuss volatility estimation. In §6.7, we discuss pricing weather derivatives contracts, including European call and put options on HDDs. In §6.8, we cover historical burn analysis as a method for pricing. In §6.9, we discuss time-series weather forecasting based on the work of Campbell and Diebold (2005). In §6.10, we give a Monte Carlo implementation in C++ to price weather options.

6.1 Weather Derivatives Market

The market for weather derivatives started in 1997 as a response from businesses effected by El Niño in order to hedge against seasonal weather risk, which can lead to significant earnings decline. The El Niño conditions were associated with warm winters in the eastern and midwestern U.S., resulting in significant energy cost savings for consumers and businesses. In addition, these conditions suppressed hurricane activities in the Atlantic and led to minimal economic losses due to hurricanes. However, the same weather pattern was also associated with extreme floods in California, resulting in both economic loss and loss of life.

After the El Niño episode, the market for weather derivatives expanded rapidly, and contracts started to be traded over-the-counter (OTC) as individually negotiated contracts. The weather derivatives market went from being essentially nonexistent in 1997 to a market in 1998 estimated at $500 million, but it was still illiquid, with large spreads and limited secondary market activity. The market grew to more than $5 billion in 2005, with better liquidity. This OTC market was primarily driven by companies in the energy sector. To increase the size of the market and to remove credit risk from the trading of the contracts, the Chicago Mercantile Exchange (CME) started an electronic marketplace for weather derivatives in September 1999. This was the first exchange where standard weather derivatives could be traded. Although weather risk has an enormous impact on many businesses, including energy producers and consumers, supermarket chains, the leisure industry, and the agricultural industries, it is primarily the energy sector that has driven the demand for weather derivatives and has caused the weather risk management industry to now evolve rapidly.[2]

The growth in the weather derivatives market in the mid-1990s can be attributed to the deregulation of the energy and utility industries in the U.S. Faced with growing competition and uncertainty in demand, energy and utility companies sought effective hedging tools to stabilize their earnings. In the deregulated environment, energy merchants quickly realized that weather conditions were the main source of revenue uncertainties. Weather affects both short-term demand and long-term supply of energy. For instance, as shown in Figure 6.1, the electricity load depends heavily on the temperature level.

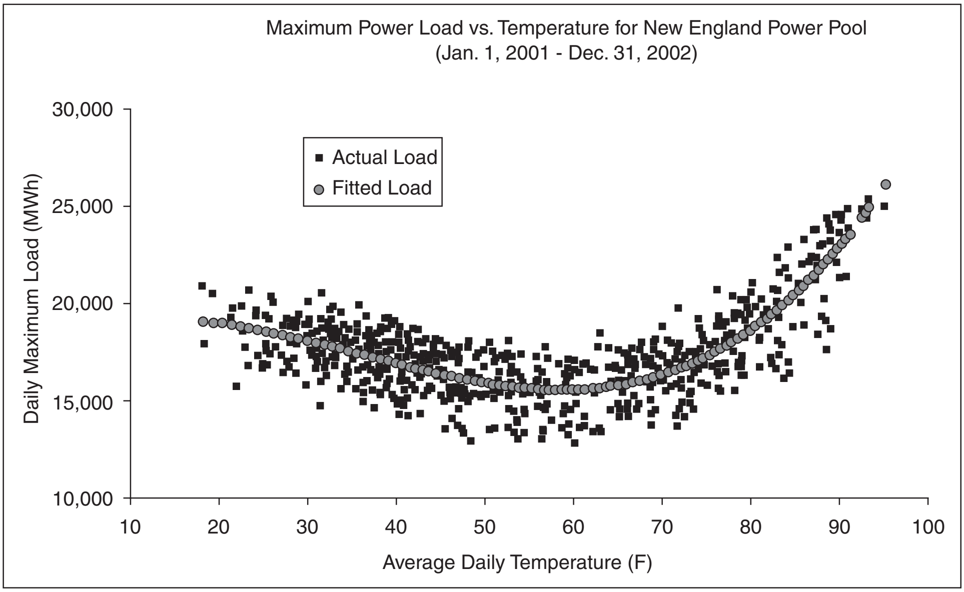

Figure 6.1 Maximum Power Load versus Temperature for New England. *Source: Cao, Li, and Wei (2004).*

The maximum power load is at the lowest when the average daily temperature is around 65°F and becomes higher when the temperature increases or decreases. Similarly, natural gas consumption is highly dependent on temperature, as shown in Figure 6.2.

Thus, short-term demand of power and energy (discussed in Chapter 7, "Energy and Power Derivatives") is largely driven by weather conditions. On the other hand, a specific pattern of weather conditions (e.g., a strong global warming trend) can also affect "the long-term supply as energy producers re-adjust their production levels."[3]

Campbell and Diebold (2005) suggest a number of interesting considerations that make weather derivatives different from "standard" derivatives. First, the underlying asset (weather) is not traded in a spot market. Second, unlike financial derivatives, which are useful for price hedging but not for quantity hedging, weather derivatives are useful for quantity hedging but not necessarily for price hedging (although the two are obviously related). That is, weather derivative products provide protection against weather-related changes in quantities, complementing extensive commodity price risk management tools already available through futures. Third, although liquidity in weather derivative markets has improved, it will likely never be as good as in traditional commodity markets, because weather is by its nature a location-specific and nonstandardized commodity, unlike, say, a specific grade of crude oil. One cannot take delivery of the weather underlying the contract. Consequently, while standard derivatives are used to hedge prices, weather derivatives are used to hedge quantity changes. For instance, a weather derivatives option might have a payoff of $1,000 for each day that it rains over the next two months in return for a premium payment of $10,000 that reflects that risk.

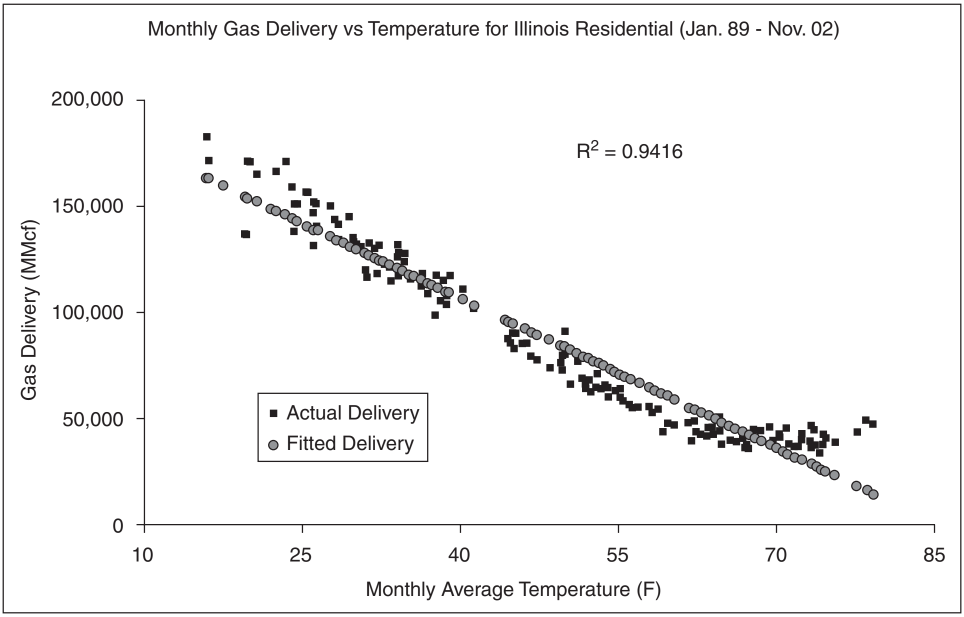

Figure 6.2 Monthly Gas Delivery versus Temperature for Illinois Residential. *Source: Cao, Li, and Wei (2004).*

The European market has not developed as quickly as the U.S. market, but there are a number of factors that suggest growth potential. One of them is the fact that Europe's energy industry is not yet fully deregulated, and as deregulation spreads throughout the industry, the volume in weather deals traded in Europe should increase. This will improve liquidity of the market and encourage new entrants into the market.

When entrants outside the energy sector become more interested in the weather derivatives market, there will also be an enormous growth potential. As mentioned earlier, there are companies in many different areas that are affected by the weather. When these companies start to look at the weather derivatives market for hedging purposes, increased liquidity as well as new products will probably follow.

Another key for the market to grow is the existence of standardized contracts. London International Financial Futures Exchange (LIFFE) is currently developing pan-European weather futures, which should increase the size of the overall weather derivatives market. There are also some barriers that must be removed if the market is to grow. For example, the quality and cost of weather data varies considerably across Europe. Companies that want to analyze their performance against historical weather data must often buy information from the national meteorological offices, and that could, in some countries, be quite expensive. It is also important that the quality of the weather data is good so that companies can rely on it when pricing derivatives.

6.2 Weather Contracts

Weather derivatives are usually structured as swaps, futures, and call/put options based on different underlying weather indices. Some commonly used indices are Heating and Cooling Degree Days, rain (precipitation), wind, stream flow, and snowfall. Many weather derivatives are based on degree-days (temperature) indices, because they are most often used. We will focus on temperature weather derivatives.

We start with some basic definitions and terminology. We define the temperature T_i as follows: Given a specific weather station, let $T_i^{\max}$ and $T_i^{\min}$ denote the maximal and minimal temperatures (in degrees Celsius) measured on day i. We define the temperature for day i as:

$$T_i = \frac{T_i^{\max} + T_i^{\min}}{2} \tag{6.1}$$

Let T_i denote the temperature for day i. Define the Heating Degree Days, HDD_i, as

$$HDD_i = \max\{65°\text{F} - T_i, 0\} \tag{6.2}$$

and the Cooling Degree Days, CDD_i, generated on that day, as

$$CDD_i = \max\{T_i - 65°\text{F}, 0\}. \tag{6.3}$$

Note that the number of HDDs or CDDs for a specific day is just the number of degrees that the temperature deviates from a reference level. It has become industry standard in the U.S. to set this reference level at 65° Fahrenheit (18°C). The names Heating and Cooling Degree Days originate from the U.S. energy sector. The reason is that if the temperature is below 18°C, people tend to use more energy to heat their homes, whereas if the temperature is above 18°C, people start turning their air conditioners on, for cooling. Most temperature-based weather derivatives are based on the accumulation of HDDs or CDDs during a certain period, usually one calendar month or a winter/summer period. Typically, the HDD season includes winter months from November to March, and the CDD season is from May to September. April and October are often referred to as the "shoulder months."

CME Weather Futures

The CME offers standardized futures contracts on temperatures based on the CME Degree Day index, which is the cumulative sum of daily HDDs or CDDs during a calendar month, as well as options on these futures. The CME Degree Day index is currently specified for 11 U.S., five European, and two Japanese cities. The HDD/CDD index futures are agreements to buy or sell the value of the HDD/CDD index at a specific future date. The notional value of one contract is $100 times the Degree Day index, and the contracts are quoted in HDD/CDD index points. The futures are cash-settled, which means that there is a daily marking-to-market based upon the index, with the gain or loss applied to the customer's account. A CME HDD or CDD call option is a contract that gives the owner the right, but not the obligation, to buy one HDD/CDD futures contract at a specific price, usually called the strike or exercise price. The HDD/CDD put option analogously gives the owner the right, but not the obligation, to sell one HDD/CDD futures contract. On the CME, the options on futures are European style, which means that they can only be exercised at the expiration date.

The CME futures have the number of HDDs (or CDDs) over one month or one season for 15 U.S. cities as the underlying temperature index. The HDD index over the time interval $[t_1, t_2]$ is defined in a continuous-time setting as $\int_{t_1}^{t_2} \max(65 - T_t, 0)dt$, whereas the CDD index is defined as $\int_{t_1}^{t_2} \max(T_t - 65, 0)dt$.[4]

For the five European cities, one can trade in futures written on the cumulative (average) temperature (CAT) index and the HDD index over a month or season. The CAT index over a timer interval $[t_1, t_2]$ is defined as $\int_{t_1}^{t_2} T_t dt$, where the temperature is measured in degrees of Celsius and not Fahrenheit. The contracts are denominated in GBP instead of USD. Moreover, the temperature level for the HDD-contracts is set at 18°C.

For the two Japanese cities (Tokyo and Osaka), the futures are written on the so-called Pacific Rim index, which measures the average daily temperature over a month or a season. The Pacific Rim index over the period $[t_1, t_2]$ is defined as $\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} T_t dt$, and the contracts are denominated in Japanese yen.[5]

The CME also offers trading in plain vanilla European options on the different temperature index futures. There exists call and put options for different strike and maturities for all HDD, CDD, CAT, and Pacific Rim index futures. Among the major market makers for the CME are Aquila Energy, Koch Energy Trading, Southern Energy, Enron, and Castlebridge Weather Markets. All these firms are also active in the OTC market for weather derivatives.

Following Benth and Šalytė–Benth (2005), consider the price dynamics of futures written on the HDD index over a specified period $[t_1, t_2]$, $t_1 < t_2$ from December to March (e.g., the winter season). Assuming a constant continuously compounding rate r, the futures price at time $t < t_1$ written on the HDD index is defined as the $\Im_t$–adapted stochastic process $F_{HDD}(\tau, t_1, t_2)$ satisfying

$$e^{-r(t_2-t)E^Q} \left[\int_{t_1}^{t_2} \max(c - T_t, 0)dt - F_{HDD}(\tau, t_1, t_2) | \Im_t \right] = 0 \tag{6.4}$$

where Q is the risk-neutral probability and c is equal to 65°F or 18°C, depending on whether the contract is for a U.S. or European city. Given the adaptedness of $F_{HDD}(t_1, t_2)$, we find the futures price to be

$$F_{HDD}(\tau, t_1, t_2) = E^Q \left[\int_{t_1}^{t_2} \max(c - T_t, 0)dt | \Im_t \right]. \tag{6.5}$$

Analogously, the CDD-futures price is

$$F_{CDD}(\tau, t_1, t_2) = E^Q \left[\int_{t_1}^{t_2} \max(T_t - c, 0) dt | \Im_t \right]. \tag{6.6}$$

By the same reasoning, one can derive the price of a CAT futures and a Pacific Rim futures to be

$$F_{CAT}(\tau, t_1, t_2) = E^Q \left[\int_{t_1}^{t_2} T_t dt | \Im_t \right], \tag{6.7}$$

and

$$F_{PRIM}(\tau, t_1, t_2) = E^Q \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} T_t dt | \Im_t \right]. \tag{6.8}$$

Note that

$$F_{PRIM}(\tau, t_1, t_2) = \frac{1}{t_2 - t_1} F_{CAT}(\tau, t_1, t_2). \tag{6.9}$$

Moreover, because $\max(c - x, 0) = c - x + \max(x - c, 0)$, we have the following:

$$F_{HDD}(\tau, t_1, t_2) = c(t_2 - t_1) - F_{CAT}(\tau, t_1, t_2) + F_{CDD}(\tau, t_1, t_2)$$

Outside the CME, there are a number of different contracts traded in the OTC market. The buyer of a HDD call, for example, pays the seller a premium at the beginning of the contract. In return, if the number of HDDs for the contract period is greater than the predetermined strike level, the buyer will receive a payoff. The size of the payoff is determined by the strike and the tick size. The *tick size* is the amount of money that the holder of the call receives for each degree-day above the strike level for the period. Often the option has a cap on the maximum payout unlike, for example, traditional options on stocks. A standard weather option can be formulated by specifying the following parameters:

- The contract type (call or put)
- The contract period (e.g., month of January)
- The underlying index (HDD or CDD)
- An official weather station from which the temperature data are obtained
- The strike level
- The tick size
- The maximum payoff (if there is any)

To find a formula for the payout of an option, let K denote the strike level and the tick size. Let the contract period consist of n days. Then, the number of HDDs and CDDs for that period are

$$H_n = \sum_{i=1}^{n} HDD_i \text{ and } C_n = \sum_{i=1}^{n} CDD_i \tag{6.10}$$

respectively. Now we can write the payoff of an uncapped HDD call as:

$$\alpha \max\{H_n - K, 0\} \tag{6.11}$$

The payouts for similar contracts like HDD puts and CDD calls/puts are defined in the same way. Figure 6.3 shows an example of HDD- and CDD-based forward and option contracts.

Examples of HDD- and CDD-Based Forward and Option Contracts

	HDD Forward	CDD Put Option
Current time	December 1, 2001	January 1, 2002
Location	Phil, Int'l Airport, Philadelphia	Hartsfield Airport, Atlanta
Long Position	ABC Bank	Air Conditioning Ltd.
Short Position	Power Supply Ltd.	XYZ Bank
Accumulation Period	February, 2002	July, 2002
Tick Size	84,000 per HDD	$10,000 per CDD
Strike Level	855 HDDs	550 CDDs
Actual Level	650 HDDs	510 CDDs
Payoffs at Maturity (Long Position)	(650 – 850) × 4000 = –$800,000	(550 – 510) × 10000 = $400,000

Figure 6.3 *Source: Cao, Li, and Wei (2004).*

6.3 Modeling Temperature

To properly model weather derivatives, we must generate the stochastic process that temperature, the underlying variable, follows and understand its behavior and movements. There are strong seasonal variations and patterns in temperature. Thus, temperature cannot be modeled well with random walks. There are other observations to consider: Temperatures exhibit high autocorrelation, which means that short-term behavior will differ from the long-term behavior. Finally, there is no underlying asset because temperature (or precipitation) cannot be bought or sold. There is no way to construct a portfolio of financial assets that replicates the payoff of a weather derivative. Given these issues with temperature, the Black Scholes framework does not apply. Figure 6.4 shows a typical pattern of average daily temperatures in Stockholm, Sweden starting January 1, 1994 and ending January 1, 2005.

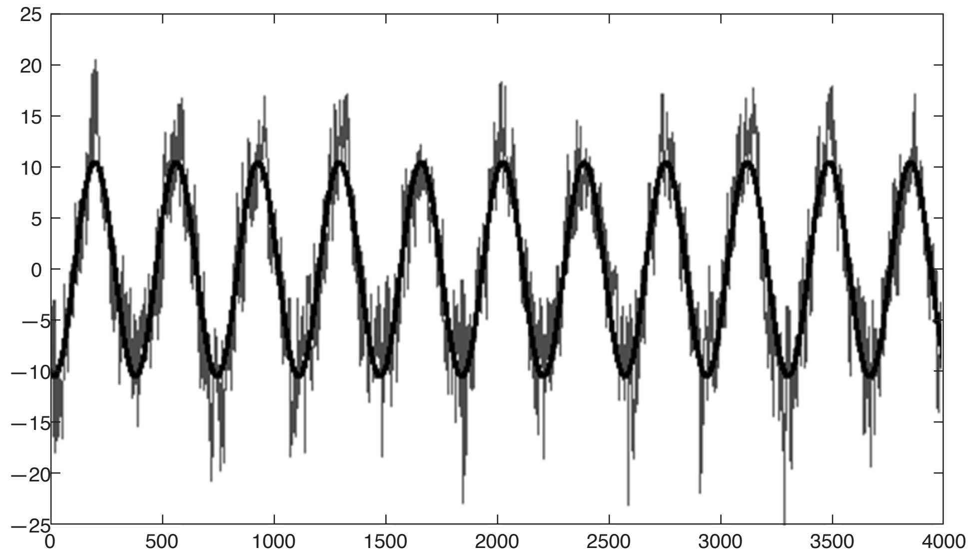

Figure 6.4 Average Temperature in Stockholm, Sweden (Jan. 2004- Jan. 2005). *Source: Alaton, P., Djehiche, B. and Sillberger, D. (2000).*

Given the seasonal and cyclical nature of temperature, the model should incorporate mean reversion in the process. The mean temperature seems to vary between about 20°C during the summers and -10°C during the winters. After a quick glance at Figure 6.1, we guess that it should be possible to model the seasonal dependence with, for example, some sine-function. This function would have the form

$$\sin(\omega t + \theta)$$

where t denotes the time, measured in days. We let $t = 1, 2, \ldots$ denote January 1, January 2, and so on. Because we know that the period of the oscillations is one year (neglecting leap years), we have $\omega = \frac{2\pi}{365}$. Because the yearly minimum and maximum mean temperatures do not usually occur at January 1 and July 1, respectively, we have to introduce a phase angle θ. Moreover, a closer look at the data series reveals a positive trend in the data. It is weak, but it does exist. The mean temperature actually increases each year. There can be many reasons to this. One reason is the fact that we may have a global warming trend all over the world. Another is the so called urban heating effect, which means that temperatures tend to rise in areas nearby a big city, because the city is growing and warming its surroundings. To catch this weak trend from data, we will assume, as a first approximation, that the warming trend is linear. We could have assumed it to be polynomial, but due to its weak effect on the overall dynamics of the mean temperature, it is only the linear term of this polynomial that will dominate.

Summing up, a deterministic model for the mean temperature at time t, T_t^m would have the form

$$T_t^m = A + Bt + C\sin(\omega t + \theta) \tag{6.12}$$

where the parameters A, B, C, and θ have to be chosen so that the curve fits the data well.

Temperature movements are not deterministic—they are affected by global and seasonal weather changes. Thus, to obtain a more realistic model, we now have to add some sort of noise to the deterministic model (6.12). One choice is a standard Wiener process, $(W_t,\ t \geq 0)$. Indeed, this is reasonable not only with regard to the mathematical tractability of the model, but also because Figure 6.5 shows a good fit of the plotted daily temperature differences with the corresponding normal distribution, although the probability of getting small differences in the daily mean temperature will be slightly underestimated.

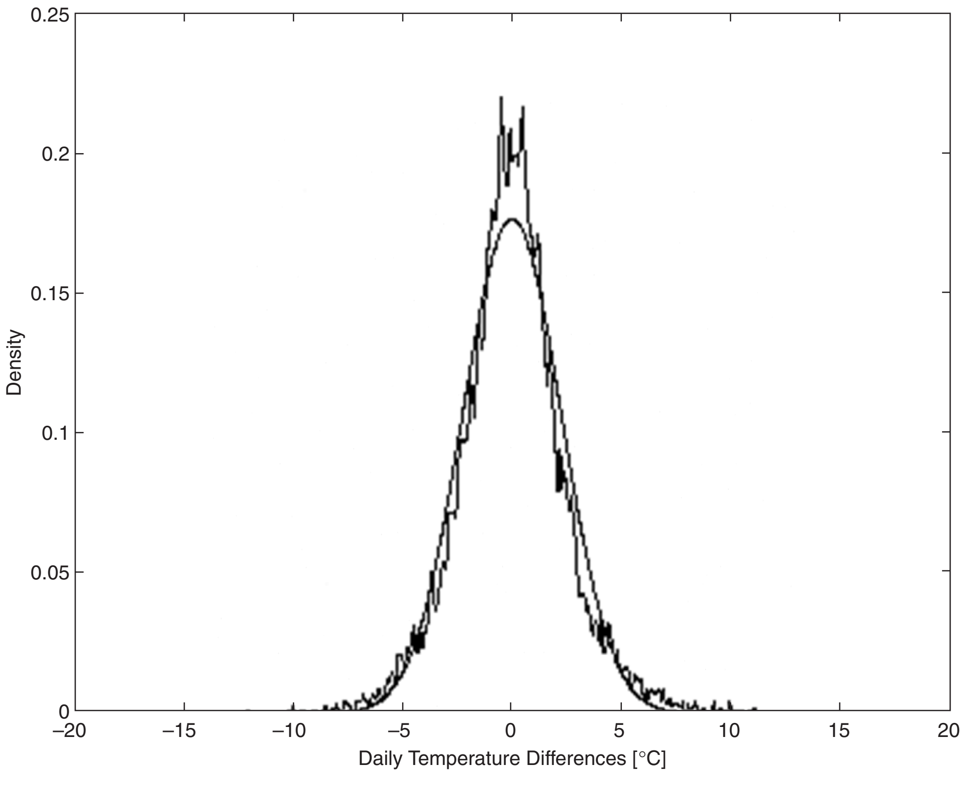

Figure 6.5 Daily Temperature Distribution. *Source: Alaton, P., Djehiche, B. and Sillberger, D. (2000), 9.*

Noise Process

A closer look at the data series reveals that the variation σ_t^2 of the temperature varies across the different months of the year, but nearly constant within each month. Especially during the winter, the quadratic variation is much higher than during the rest of the year. Therefore, we make the assumption that σ_t is a piecewise constant function, with a constant value during each month.[6] We specify σ_t as

$$\sigma_t = \begin{cases} \sigma_1 \text{ during January} \\ \sigma_2 \text{ during February} \\ \vdots \\ \sigma_n \text{ during December} \end{cases}$$

where $\sigma_i, i = 1, ..., 12$ are positive constants. Thus, the driving noise process of the temperature would be $(\sigma_t W_t, t \geq 0)$.

Mean-Reversion

We know that the temperature cannot, for example, rise day after day for a long time. This means that our model should not allow the temperature to deviate from its mean value for more than short periods of time. In other words, the stochastic process describing the temperature we are looking for should have a mean-reverting property. Putting all the assumptions together, we model temperature by a stochastic process solution of the following SDE[7]

$$dT_t = a(T_t^m - T_t)dt + \sigma_t dW_t \tag{6.13}$$

where $a \in \Re$ determines the speed of mean-reversion. The solution of such an equation is usually called an Ornstein-Uhlenbeck process. Alaton et al. (2000) suggests adding another term to the drift because equation (6.13) is not actually mean-reverting to T_t^m in the long run, as shown by Dornier and Queruel (2000):[8]

$$\frac{dT_t^m}{dt} = B + \omega C \cos(\omega t + \theta) \tag{6.14}$$

Given the mean temperature T_t^m is not constant, this term will adjust the drift so that the solution of the SDE has the long run mean T_t^m.

Starting at $T_s = x$, we now get the following model for the temperature

$$dT_t = \left(\frac{dT_t^m}{dt} + a(T_t^m - T_t) \right) dt + \sigma_t dW_t, \ t>s \tag{6.15}$$

whose solution is

$$T_t = (x - T_s^m)e^{-a(t-s)} + T_t^m + \int_s^t e^{-a(t-\tau)} \sigma_\tau dW_\tau \tag{6.16}$$

where

$$T_t^m = A + Bt + C \sin(\omega t + \theta). \tag{6.17}$$

6.4 Parameter Estimation

To estimate the unknown parameters A, B, C, θ, a, and σ, one can use least squares regression techniques. To find the numerical values of the parameters in (6.17), we can estimate the parameters of

$$Y_t = a_1 + a_2 t + a_3 \sin(\omega t) + a_4 \cos(\omega t) \tag{6.18}$$

by fitting the function in (6.18) to the temperature data using the method of least squares. We must find the parameter vector $\xi = (a_1, a_2, a_3, a_4)$ that solves

$$\min_{\xi} ||\mathbf{Y} - \mathbf{X}||^2 \tag{6.19}$$

where $\mathbf{Y}$ is the vector with elements (6.18) and $\mathbf{X}$ is the (historical) data vector. The constant in the model (6.17) are then obtained by

$$A = a_1 \tag{6.20}$$

$$B = a_2 \tag{6.21}$$

$$C = \sqrt{a_3^2 + a_4^2} \tag{6.22}$$

$$\theta = \arctan\left(\frac{a_4}{a_3}\right) - \pi \tag{6.23}$$

As an example, Alaton et al. (2000) estimates the parameters in (6.18) using temperature data from the Bromma Airport in Stockholm, Sweden from the last 40 years. The following function for the mean temperature is given by:

$$T_t^m = 5.97 + 6.57 \cdot 10^{-5} t + 10.4 \sin\left(\frac{2\pi}{365} t - 2.01\right) \tag{6.24}$$

The amplitude of the sin-function is about 10°C, which means that the difference in temperature between a typical winter day and a summer day is about 20°C.[9]

6.5 Volatility Estimation

Alaton et al. (2000) provides to estimators for σ from the data (collected for each month). Given a specific month μ of N_i days, denote the outcomes of the observed temperatures during the month μ by $T_j, j = 1, \ldots, N_i$. The first estimator is based on the quadratic variation of T_t (see Basawa and Prasaka Rao [1980], 212–213).[10]

$$\widehat{\sigma}_i^2 = \frac{1}{N} \sum_{j=0}^{N_\mu - 1} (T_{j+1} - T_j)^2 \tag{6.25}$$

The second estimator is derived by discretizing (6.15) and viewing the discretized equation as a regression equation. During a given month μ, the discretized equation is

$$T_j = T_j^m - T_{j-1}^m + a T_{j-1}^m + (1-a) T_{j-1} + \sigma_\mu \varepsilon_{j-1},\ j = 1, \ldots, N_\mu \tag{6.26}$$

where $\{\varepsilon_j\}_{j=1}^{N_m-1}$ are independent standard normally distributed random variables—e.g., $\varepsilon_j \sim N(0,1)$. Let $\widetilde{T}_j = T_j - (T_j^m - T_{j-1}^m)$. Then we can write (6.26) as

$$\widetilde{T}_j = aT_{j-1}^m + (1-a)T_{j-1} + \sigma_\mu \varepsilon_{j-1}, \tag{6.27}$$

which can be seen as a regression of today's temperature on yesterday's temperature. Thus, an efficient estimator of σ_i is (see Brockwell and Davis [1990]):[11]

$$\widehat{\sigma}_\mu^2 = \frac{1}{N_\mu - 2}\sum_{j=0}^{N_\mu}(\widetilde{T}_j - aT_{j-1}^m + (1-a)T_{j-1})^2 \tag{6.28}$$

To evaluate the estimator in (6.28), one needs to find an estimator of the mean-reversion parameter a, which is done in the next section.

6.6 Mean-Reversion Parameter Estimation

The mean-reversion parameter can be estimated using the martingale estimation function methods given by Bibby and Sorensen (1995).[12] Let $b(T_t; a)$ denote the drift term of the temperature stochastic process in (6.15):

$$b(T_t; a) = \frac{dT_t^m}{dt} + a(T_t^m - T_t) \tag{6.29}$$

Following Alaton et al. (2000), based on observations collected during n days, an efficient estimator $\widehat{a}_n$ of a is obtained as a zero of equation:

$$G_n(\widehat{a}_n) = 0 \tag{6.30}$$

where

$$G_n(a) = \sum_{i=1}^{n} \frac{\dot{b}(T_{i-1}; a)}{\sigma_{i-1}^2}(T_i - E[T_i | T_{i-1}]) \tag{6.31}$$

and $\dot{b}(T_t; a)$ denotes the derivatives of the drift in (6.29) with respect to a. To solve (6.30), one only needs to determine each of the conditional expectation terms $E[T_i|T_{i-1}]$ in (6.31). By equation (6.16) for $t \geq s$,

$$T_t = (T_s - T_s^m)e^{-a(t-s)} + T_t^m + \int_s^t e^{-a(t-\tau)}\sigma_\tau dW_\tau \tag{6.32}$$

which yields

$$E[T_i|T_{i-1}] = (T_{i-1} - T_{i-1}^m)e^{-a} + T_i^m \tag{6.33}$$

where as before

$$T_t^m = A + Bt + C\sin(\omega t + \theta).$$

Therefore,

$$G_n(a) = \sum_{i=1}^{n} \frac{T_{i-1}^m - T_{i-1}}{\sigma_{i-1}^2} \left(T_i - (T_{i-1} - T_{i-1}^m)e^{-a} - T_i^m \right) \tag{6.34}$$

from which one finds the mean-reversion estimator

$$\widehat{a}_n = -\log \left(\frac{\sum_{i=1}^{n} Y_{i-1}(T_i - T_i^m)}{\sum_{i=1}^{n} Y_{i-1}(T_{i-1} - T_{i-1}^m)} \right) \tag{6.35}$$

is the unique zero of equation (6.30), where

$$Y_{i-1} = \frac{T_{i-1}^m - T_{i-1}}{\sigma_{i-1}^2} \quad i = 1, 2, \ldots, n. \tag{6.36}$$

6.7 Pricing Weather Derivatives

Model Framework

We define the following parameters for the pricing model:

- T_t^m: modeled average temperature at time t
- T_t: current temperature at time t
- a: mean reversion parameter
- σ: volatility of the temperature
- W_t: Wiener process

The market for weather derivatives is a typical example of an incomplete market, because the underlying variable, the temperature, is not tradable. Therefore, we have to consider the market price of risk λ, in order to obtain unique prices for such contracts. Because there is not yet a real market from which we can obtain prices, we assume for simplicity that the market price of risk is constant. Furthermore, we assume that we are given a risk-free asset with constant interest rate r and a contract that for each degree Celsius pays one unit of currency. Thus, under a martingale measure Q, characterized by the market price of risk λ, the temperature process also denoted by T_t satisfies the following dynamics:

$$dT_t = \left(\frac{dT_t^m}{dt} + a(T_t^m - T_t) - \lambda\sigma_t \right) dt + \sigma_t dW_t \tag{6.37}$$

where $\{W_t, t \geq 0\}$ is a Q–Wiener process. Following Alaton et al. (2000), we start computing the expected value and the variance of T_t because the price of a derivative is expressed as a discounted expected value under the risk-neutral martingale measure Q. We will use Girsanov's theorem to change to the drift under the physical measure P. However, because

Girsanov's transformation only changes the drift, the variance of T_t is the same under both measures. Therefore,

$$\mathrm{Var}[T_t|\Im_s] = \int_s^t \sigma_u^2 e^{-2a(t-u)} du. \tag{6.38}$$

Furthermore, it follows from (6.16) that

$$E^P[T_t|\Im_s] = (T_s - T_s^m)e^{-a(t-s)} + T_t^m. \tag{6.39}$$

Thus, in view of equation (6.37), we must have the following:

$$E^Q[T_t|\Im_s] = E^P[T_t|\Im_s] - \int_s^t \lambda\sigma_u e^{-a(t-u)} \tag{6.40}$$

Evaluating the integrals in one of the intervals where σ is constant, we get that

$$E^Q[T_t|\Im_s] = E^P[T_t|\Im_s] - \frac{\lambda\sigma_i}{a}(1 - e^{-a(t-s)}) \tag{6.41}$$

and the variance is

$$\mathrm{Var}[T_t|\Im_s] = \frac{\sigma^2}{2a}\left(1 - e^{-2a(t-s)}\right). \tag{6.42}$$

The covariance of the temperature between two different days is for $0 \leq s \leq t \leq u$:

$$\mathrm{Cov}[T_t, T_u|\Im_s] = e^{-a(u-t)}\,\mathrm{Var}[T_t|\Im_s] \tag{6.43}$$

As Alaton et al. (2000) shows, suppose now that t_1 and t_n denote the first and last day of a month and start the process at some time s from the month before $[t_1, t_n]$. To compute the expected value and variance of T_t in this case, we split the integrals in (6.40) and (6.38) into two integrals where σ is constant in each one. We then get

$$E^Q[T_t|\Im_s] = E^P[T_t|\Im_s] - \frac{\lambda}{a}(\sigma_i - \sigma_j)e^{-a(t-t_1)} + \frac{\lambda\sigma_i}{a}e^{-a(t-s)} - \frac{\lambda\sigma_j}{a} \tag{6.44}$$

and the variance is

$$\mathrm{Var}[T_t|\Im_s] = \frac{1}{2a}\left(\sigma_i^2 - \sigma_j^2\right)e^{-2a(t-t_1)} - \frac{\sigma_i^2}{2a}e^{-2a(t-s)} + \frac{\sigma_j^2}{2a}. \tag{6.45}$$

The generalization to a larger time interval now becomes straightforward.

Pricing a Heating Degree Day Option

To price a standard HDD call option, we use the payoff given in (6.5)

$$\chi = \alpha \max(H_n - K, 0) \tag{6.46}$$

where

$$H_n = \sum_{i=1}^{n} \max(65^\circ F - T_t, 0). \tag{6.47}$$

The payoff depends on the accumulation of HDDs during some time period (e.g., in the winter during the month of January).

For simplicity, we assume that the tick size is $\alpha = 1$. The contract in (6.46) is a type of arithmetic average Asian option. In the case of a log-normally distributed underlying process, no exact formula for the price of such an option is known, so we try to make some sort of approximation. We know that under the risk-neutral measure Q, and given information at time s,

$$T_t \sim N(\mu_t, \sigma_t^2)$$

where μ_t is given by (6.44) and σ_t^2 is given by (6.45). Suppose the probability that max $(65 - T_t, 0) = 0$ is extremely small on a winter day. Then, for such a contract, we can write

$$H_n = 65n - \sum_{i=1}^{n} T_{t_i}. \tag{6.48}$$

We now determine fair pricing under a (normal) distributional model. Following Alaton et al. (2000), we know that $T_t, i = 1, ..., n$ are all samples from an Ornstein-Uhlenbeck process, which is a Gaussian process. This means that the vector $(T_{t_1}, T_{t_2}, \ldots, T_{t_n})$ is Gaussian. Because the sum of (6.48) is a linear combination of the elements of the vector, H_n is also Gaussian.[13] We compute the first and second moments. We have, for $t < t_1$,

$$E^Q[H_n|\Im_s] = E^Q\left[65n - \sum_{i=1}^{n} T_{t_i}|\Im_s\right] = 65n - \sum_{i=1}^{n} E^Q[T_{t_i}|\Im_t] \tag{6.49}$$

and

$$\mathrm{Var}[H_n|\Im_s] = \sum_{i=1}^{n} \mathrm{Var}[T_{t_i}|\Im_s] + 2\sum_{j=1}^{n}\sum_{i<j} \mathrm{Cov}[T_{t_i}, T_{t_j}|\Im_t]. \tag{6.50}$$

Suppose we make the previous calculations and find that:

$$E^Q[H_n|\Im_t] = \mu_n \text{ and } \mathrm{Var}[H_n|\Im_t] = \sigma_n^2 \tag{6.51}$$

Then, $H_n \sim N(\mu_n, \sigma_n)$. Thus, the price at $t \le t_1$ of the claim (6.46) is given by

$$\begin{aligned} c_{HDD}(t) &= e^{-r(t_n - t)} E^Q[\max(H_n - K, 0)|\Im_t] \\ &= e^{-r(t_n - t)} \int_K^\infty (x - K) f_{H_n}(x) dk \\ &= e^{-r(t_n - t)} \left((\mu_n - K)\Phi(-\alpha_n) + \frac{\sigma_n}{\sqrt{2\pi}} e^{-\frac{\alpha_n^2}{2}} \right) \end{aligned} \tag{6.52}$$

where $\alpha_n = (K - \mu_n)/\sigma_n$ and Φ denotes the cumulative standard normal distribution function (see also Platen and West [2004], pg. 23[14]).

Similarly, we can derive the formula for the price of an HDD put option with the claim payoff of:

$$y = \max(K - H_n, 0) \tag{6.53}$$

The price is

$$\begin{aligned} p_{HDD}(t) &= e^{-r(t_n - t)} E^Q[\max(K - H_n, 0)|\Im_t] \\ &= e^{-r(t_n - t)} \int_0^K (K - x) f_{H_n}(x) dx \\ &= e^{-r(t_n - t)} \left\{ (K - \mu_n) \left(\Phi(\alpha_n) - \Phi\left(-\frac{\mu_n}{\sigma_n} \right) \right) + \right. \\ &\quad \left. \frac{\sigma_n}{\sqrt{2\pi}} \left(e^{-\frac{\alpha_n^2}{2}} - e^{-\frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2} \right) \right\}. \end{aligned} \tag{6.54}$$

Formulas (6.52) and (6.54) for call and put, respectively, hold primarily for contracts during winter months, which typically is the period from November to March.[15] During the summer months, we cannot use these formulas without restrictions. If the mean temperatures are very close to, or even higher than, 65°F (18°C), we no longer have $\max(65 - T_{t_i}, 0) \neq 0$. For such contracts, we can use Monte Carlo simulation.

Zeng (2000) discusses a pure actuarial statistical approach (see also Platen and West [2004]), as well as a prediction-based pricing approach for weather derivatives. In the latter approach, a normal distribution is fit to the historical CDD data by assuming that the option CDD follows a normal distribution with the mean and standard deviation equal to the sample mean and standard deviation, respectively, of the historical data. The probability of "non-exceedance" is evenly divided into the highest, middle, and lowest thirds, as shown in Figure 6.6.

The fitted distribution is then sampled such that the number of samples corresponding to the highest, middle, and lowest thirds are proportional to the above, near, and below climate norms (denoted p_A, p_N, and p_B, respectively). For instance, unlike traditional Monte Carlo, which samples the fitted distribution evenly across the probabilistic distribution, the prediction-based approach only samples from the highest, middle, and lowest thirds, respectively, of the distribution so as to incorporate the probabilistic climate prediction into the sample CDD values (a biased sampling Monte Carlo approach). The payoffs are averaged over from this sampled distribution just like in traditional Monte Carlo. The method exploits the fact that the rank-order correlation of the historical data will be high given the seasonal nature of temperature (for instance, a high July temperature tends to be associated with a high June-July-August (JJA) temperature), and that the predicted anomaly probabilities for the temperature p_A, p_N, and p_B are assumed to approximate the probabilities that the CDD will be above, near, and below the climate norm, respectively.[16]

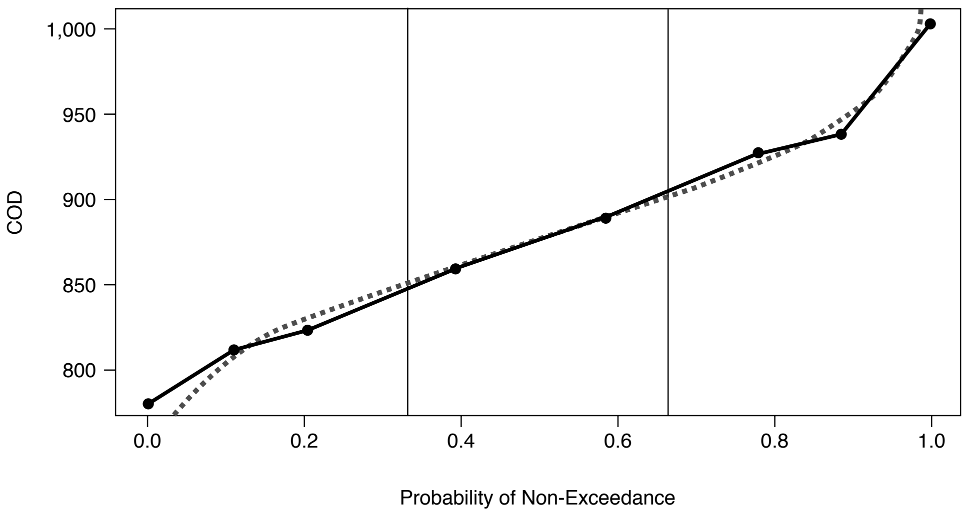

Figure 6.6 The probability of non-exceedance. *Source: Zeng (2000).*

6.8 Historical Burn Analysis

The method of historical burn analysis evaluates the contract against historical data and takes the average of realized payoffs as the fair value estimate. For instance, suppose a call option is written on a city's CDDs for the month of July, and suppose we have 20 years of daily temperatures. To apply the historical burn analysis method, for each July of the past 20 years, we calculate the option payoff using the realized CDDs. The average of the 20 payoffs is the estimate for the call option value. Thus, this method's key assumption is that the past always reflects the future on average. To be more precise, the method assumes that the distribution of the past payoffs accurately depicts the future payoff's distribution. This is a far-reaching requirement in most cases. For instance, we have only 20 payoff observations in the previous example, which can hardly capture the complete characteristics of the true distribution.

Cao, Li, and Wei (2004) apply this method to call options written on the three-month (January, February, and March) cumulative HDDs for Atlanta, Chicago, and New York. Table 6.1 shows the calculations. They first calculate the realized cumulative HDDs for each year, and then evaluate the option's payoff[17] accordingly. (The exercise prices are set at 1,500, 3,200, and 2,500 for Atlanta, Chicago, and New York, respectively.) When Cao, Li, and Wei (2004) use all 20 historical observations for Atlanta, for example, the average payoff is 92.15; when the most recent 19 observations are used, the average payoff is 82.37, and so on. When they use only the most recent 10 observations, the fair value estimate is 22.50. Going back 10 years versus 20 years would lead to a difference of more than 300% in value estimates! The estimates for the option on New York's HDDs have the smallest dispersion. But even there, the highest estimate is 70% higher than the lowest estimate.

We can argue that we should use as long a time series as possible to enhance accuracy. However, using more data will cover more temperature variations, but a derivative

Table 6.1

	HDDs for Jan. 1 - March 31			HDD Call Option Payoff			Years in	Estimate of Call Option Value		
Year	Atlanta	Chicago	New York	Atlanta	Chicago	New York	Average	Atlanta	Chicago	New York
1979	1778	3851	2841	278	651	341				
1980	1672	3474	2702	172	274	202				
1981	1698	3183	2708	198	0	208				
1982	1587	3760	2791	87	560	291				
1983	1749	3002	2443	249	0	0				
1984	1660	3424	2724	160	224	224				
1985	1723	3591	2567	223	391	67				
1986	1416	3208	2533	0	8	33				
1987	1602	2813	2504	102	0	4				
1988	1649	3417	2593	149	217	93	10	22.50	63.70	69.60
1989	1242	3150	2417	0	0	0	11	34.00	77.64	71.73
1990	1009	2627	2078	0	0	0	12	39.67	71.17	66.08
1991	1354	3066	2217	0	0	0	13	36.62	66.31	63.54
1992	1325	2862	2439	0	0	0	14	49.93	89.50	63.79
1993	1514	3277	2667	14	77	167	15	57.27	98.47	74.47
1994	1410	3524	2921	0	324	421	16	69.25	92.31	69.81
1995	1295	3096	2370	0	0	0	17	70.29	119.82	82.82
1996	1666	3410	2608	166	210	108	18	77.39	113.17	89.78
1997	1102	3226	2377	0	26	0	19	82.37	121.63	95.68
1998	1545	2637	2060	45	0	0	20	92.15	148.10	107.95
							Highest	92.15	148.10	107.95
							Lowest	22.50	63.70	63.54
							Highest/Lowest	4.10	2.32	1.70

Source: Cao, Li, and Wei (2004)

security's payoff depends on the future temperature behavior, which may be quite different from past history. This is especially true if the maturity of the derivative security is short. Ultimately, the decision boils down to a trade-off between statistical power and representativeness.[18] The commonly accepted sample length in the industry appears to be between 20 to 30 years. Furthermore, we could combine the burn analysis with temperature forecasts to arrive at a more representative price estimate.

Like an insurance or actuarial method, the method of historical burn analysis is incapable of accounting for the market price of risk associated with the temperature variable.[19] These methods are only useful from the perspective of a single dealer.[20] To establish a unique market price that incorporates a risk premium, we need a dynamic, forward-looking model such as

$$dY(t) = \beta[\theta(t) - Y(t)]dt + \sigma(t)Y_t^r dW(t) \tag{6.55}$$

where $Y(t)$ is the current temperature, $\theta(t)$ is the deterministic long-run level of the temperature, β is the speed at which the instantaneous temperature reverts to the long-run level $\theta(t)$, $\sigma(t)$ is the volatility (which is season-dependent), $r = 0$, 0.5, or 1, and $W(t)$ is a Wiener process that models the temperature's random innovations. The process in (6.55) needs to be discretized as

$$Y_t - Y_{t-1} = \beta[\theta(t) - Y_t]\Delta t + \sigma(t)Y_t^r \Delta W_t$$

in order to estimate β and the parameters imbedded in $\theta(t)$ and $\sigma(t)$, where the functional forms for $\theta(t)$ and $\sigma(t)$ can be specified based on careful statistical analyses. Once the process in (6.55) is estimated using a method like ordinary least squares, we can then value any contingent claim by taking expectation of the discounted future payoff; i.e.,

$$X = e^{-r(T-t)} E\left[g(Y_t, Y_{t+1}, ..., Y_T)\right] \tag{6.56}$$

where X is the current value of the contingent claim, r is the risk-free interest rate, T is the maturity of the claim, and $g(Y_t, Y_{t+1}, ..., Y_T)$ is the payoff at time T, which usually is a function of the realized temperatures, $Y_t, Y_{t+1}, ..., Y_T$ (e.g., a contract on cumulative HDDs or CDDs).

Given the complex form of $\theta(t)$ and $\sigma(t)$ and the path-dependent nature of most payoffs, the formula in (6.55) usually does not have closed-form solutions. Monte Carlo simulations must be used. There are two additional drawbacks of this continuous setup. First, it does not allow a place for the market price of risk. Instead, a risk-neutral valuation is imposed without any theoretical justification. Second, the process in (6.55) cannot reflect the persistent serial correlations typically present in daily temperatures.[21] As a result of these drawbacks, time-series discrete processes and models that can incorporate serial correlation have been proposed for forecasting the temperature, as discussed in the next section (see Campbell and Diebold [2002]; Cao and Wei [2003]).

6.9 TIME-SERIES WEATHER FORECASTING

As Figure 6.4 shows, the daily average temperature moves in a given city repeatedly and regularly through periods of high temperature (summer) and low temperature (winter). However, seasonal fluctuations differ noticeably across cities both in terms of amplitude and detail of pattern.[22] But most cities' unconditional temperature distributions are bimodal, with peaks characterized by cool and warm temperatures.[23] Temperature time-series across cities suggests that a seasonal component will be an important factor in any time-series model fit to daily average temperature, as "average temperature displays pronounced seasonal variation, with both the amplitude and precise seasonal patterns differing noticeably across cities."[24] Campbell and Diebold (2000) use a low-ordered Fourier series as opposed to daily dummies to model seasonality for two reasons. First, use of a low-ordered Fourier series produces a smooth seasonal pattern, which accords with the basic intuition that "the progression through different seasons is gradual rather than discontinuous."[25] Second, the Fourier approximation "produces a huge reduction in the number of parameters to be estimated, which significantly reduces computing time and enhances numerical stability."[26]

Campbell and Diebold (2000) also incorporate nonseasonal factors that may be operative in the dynamics of daily average temperature, though dominated by seasonality; in particular, a deterministic linear trend and cycle—persistent (but covariance stationary) dynamics apart from trend and seasonality.[27] These cyclical dynamics are captured using autoregressive lags, which facilitates numerically stable parameter estimates. Thus, an autoregressive model to forecast or estimate future average temperature is suggested:

$$T_t = \mathit{Trend}_t + \mathit{Seasonal}_t + \sum_{l=1}^{L} \rho_{t-1} T_{t-l} + \sigma \varepsilon_t \tag{6.57}$$

where

$$\mathit{Trend}_t = \beta_0 + \beta_1 t$$

and

$$\mathit{Seasonal}_t = \sum_{p=1}^{P} \left(\delta_{c,p} \cos\left(2\pi p \frac{d(t)}{365}\right) + \delta_{s,p} \sin\left(2\pi p \frac{d(t)}{365}\right)\right) \tag{6.58}$$

$$\varepsilon_t \overset{iid}{\sim} N(0,1)$$

and where $d(t)$ is a repeating step function that cycles through $1, \ldots, 365$ (i.e., each day of the year assumes one value between 1 and 365). The model is estimated using ordinary least squares, regressing average temperature on constant, trend, Fourier, and lagged average temperature terms, using $L = 25$ autoregressive lags (in order to capture long-memory dynamics) and three Fourier sine and cosine terms ($P = 3$).

Campbell and Diebold (2002) find that the model in (6.57), based on the correlograms of the *squared* residuals, has conditional heteroskedasticity in the model and has drastic misspecification related to nonlinear dependence despite residual autocorrelations that are negligible and consistent with white noise. As a result, Campbell and Diebold (2002) respecify the model as

$$T_t = \mathit{Trend}_t + \mathit{Seasonal}_t + \sum_{l=1}^{L} \rho_{t-1} T_{t-l} + \sigma_t \varepsilon_t \tag{6.59}$$

where

$$Trend_t = \beta_0 + \beta_1 t$$

and

$$Seasonal_t = \sum_{p=1}^{P} \left(\delta_{c,p} \cos\left(2\pi p \frac{d(t)}{365}\right) + \delta_{s,p} \sin\left(2\pi p \frac{d(t)}{365}\right)\right) \tag{6.60}$$

$$\sigma_t^2 = \sum_{q=1}^{Q} \left(\gamma_{c,q} \cos\left(2\pi q \frac{d(t)}{365}\right) + \gamma_{s,q} \sin\left(2\pi q \frac{d(t)}{365}\right)\right) + \sum_{r=1}^{R} \alpha_r \varepsilon_{t-r}^2 \tag{6.61}$$

$$\varepsilon_t \overset{iid}{\sim} N(0,1)$$

where, as before, $d(t)$ is a repeating step function that cycles through 1, . . . ,365 (i.e., each day of the year assumes one value between 1 and 365), and one sets $L = 25$, $P = 3$, $Q = 2$, and $R = 1$.

Model (6.57) is identical to model (6.59) with the addition of the conditional variance equation (6.61), which allows for two types of volatility dynamics. First, it captures seasonality volatility by approximating seasonality in the conditional variance in the same way as equation (6.60) approximates seasonality in the conditional mean, via a Fourier series.[28] Second, the variance equation captures "autoregressive effects in the conditional variance movements, which often arise naturally in time-series contexts, in which shocks to the conditional variance may have effects that persist for several periods, precisely as in the seminal work of Engle (1982)."[29] The model is estimated using Engle's (1982) asymptotically efficient two-step approach. First, equation (6.59) is estimated by ordinary least squares, regressing average temperature on constant, trend, Fourier, and lagged average temperature terms. Second, the variance equation (6.61) is estimated by regressing squared residuals from equation (6.59) on constant, Fourier, and lagged squared residual terms. Square roots of the inverse of fitted values $\widehat{\sigma}_t^{-1}$ are used as weights in a weighted least squares re-estimation of (6.59).[30]

Campbell and Diebold (2002) show that the estimated model for conditional heteroskedasticity reduces (but does not eliminate) residual excess kurtosis. Figures 6.7 and 6.8 show the correlograms and correlograms of squared residuals, respectively, for the daily average temperature in various U.S. cities.

Each panel displays autocorrelations of the squared residuals from our daily average temperature model

$$\left(T_t - Trend_t + Seasonal_t + T_t - Trend_t + Seasonal_t + \sum_{l=1}^{L} \rho_{t-1} T_{t-l} \right)^2$$

together with approximate 95% confidence intervals under the null hypothesis of white noise.[31] The correlograms shows that there was no evidence of serial correlation in the standardized residuals. The correlograms of the squared standardized residuals from model (6.59) are substantial improvement over those from model (6.55), suggesting that there is no significant deviation from white noise behavior and that the fitted model (6.59) is adequate.[32]

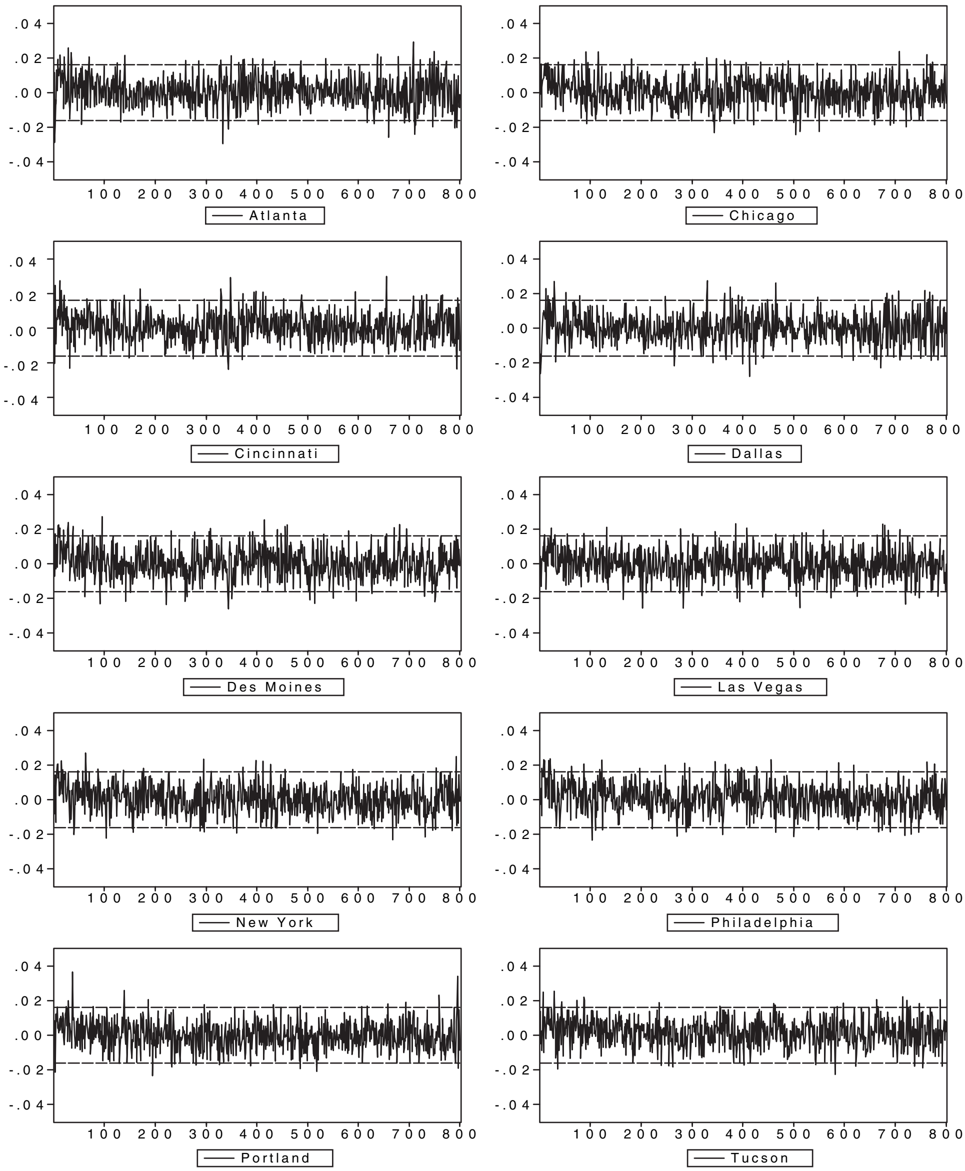

Figure 6.7 Correlograms. *Source: Campbell and Diebold (2002).*

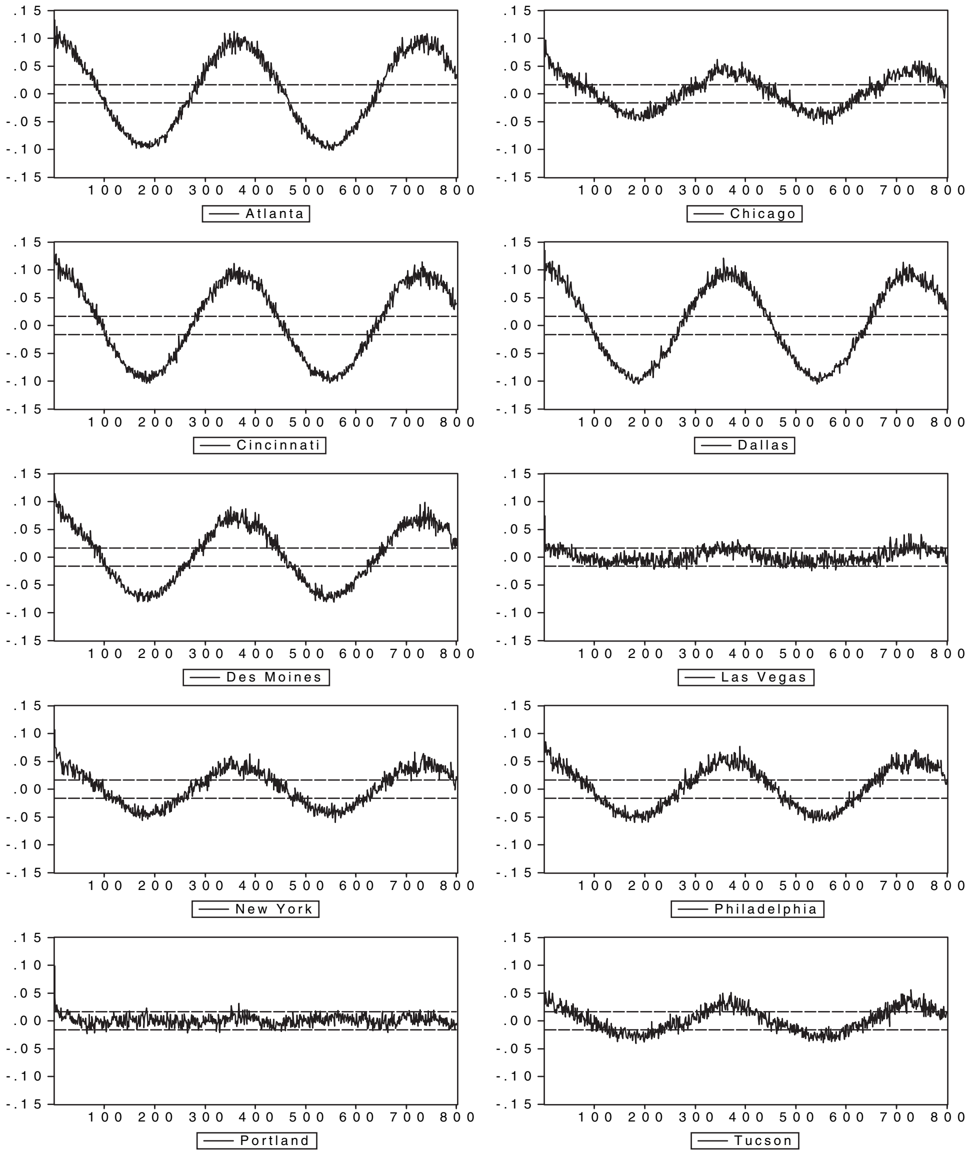

Figure 6.8 Correlograms of squared residuals. *Source: Campbell and Diebold (2002).*

Figure 6.9 shows the actual values, fitted values, and residuals of the daily average temperature in various cities in the U.S.

Each panel displays actual values, fitted values, and residuals from an unobserved-components model

$$T_t = \mathit{Trend}_t + \mathit{Seasonal}_t + \sum_{l=1}^{L} \rho_{t-1} T_{t-l} + \sigma_t \varepsilon_t.$$

Figure 6.10 shows the estimated conditional standard deviations of the daily average temperature in various U.S. cities.

Each panel displays a time series of estimated conditional standard deviations of daily average temperature obtained from the model:

$$\widehat{\sigma}_t = \sum_{q=1}^{2} \left(\widehat{\gamma}_{c,q} \cos \left(2\pi q \frac{d(t)}{365} \right) + \widehat{\gamma}_{s,q} \sin \left(2\pi q \frac{d(t)}{365} \right) \right) + \widehat{\alpha} \varepsilon_{t-1}^2$$

Figure 6.11 shows the estimated seasonal patterns of daily average temperature (Fourier series versus daily dummies).

The panels shows smooth seasonal patterns estimated from Fourier models,

$$\mathit{Seasonal}_t = \sum_{p=1}^{3} \left(\delta_{c,p} \cos \left(2\pi p \frac{d(t)}{365} \right) + \delta_{s,p} \sin \left(2\pi p \frac{d(t)}{365} \right) \right),$$

and rough seasonal patterns estimated from dummy variable models,

$$\mathit{Seasonal}_t = \sum_{t=1}^{365} \delta_i D_{it}.$$

Figure 6.12 shows histograms of the estimated unconditional distributions of daily average temperature in various U.S. cities from 1996–2001.

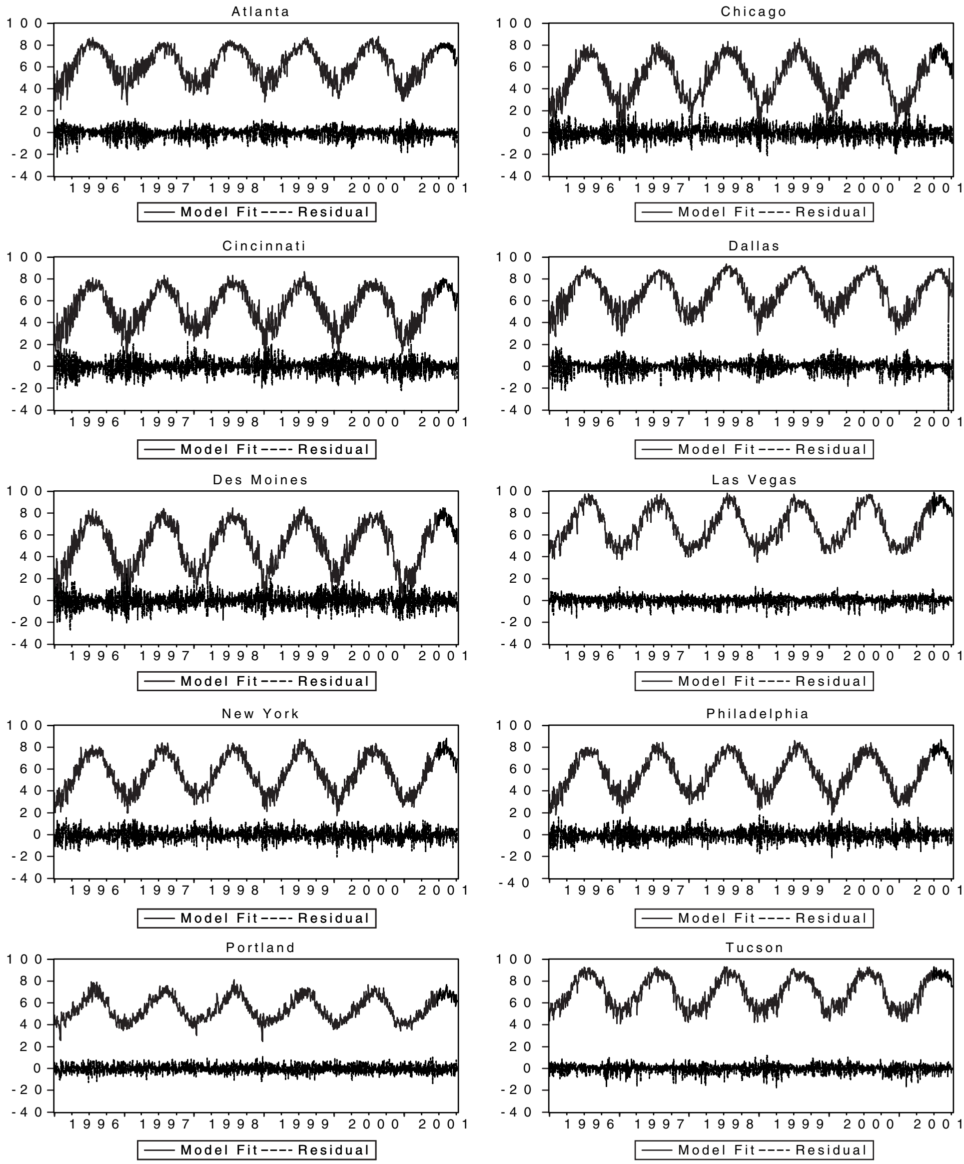

Figure 6.9 Actual values, fitted values, and residuals of the daily average temperature in various cities in the U.S. *Source: Campbell and Diebold (2002).*

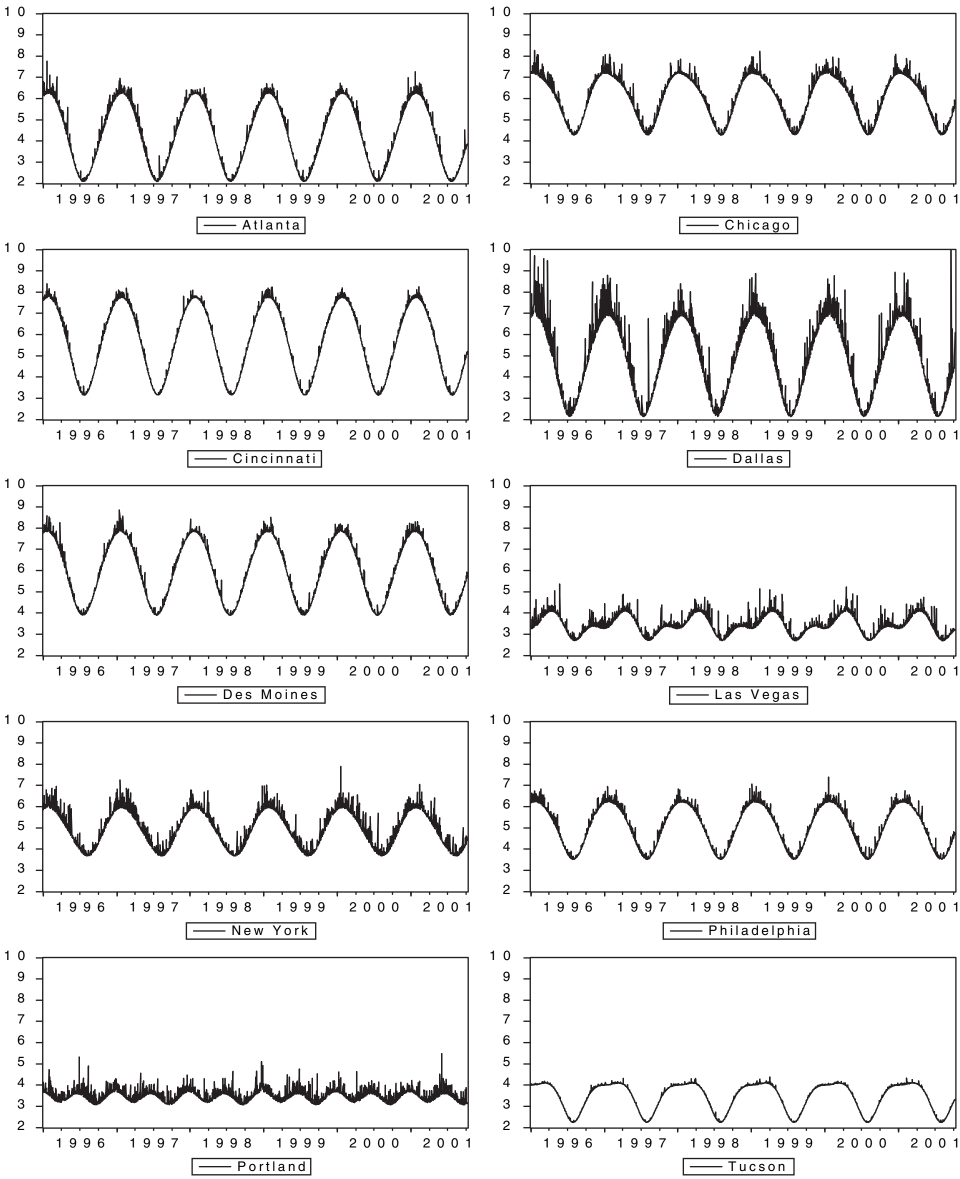

Figure 6.10 Estimated conditional standard deviations of the daily average temperature in various U.S. cities. *Source: Campbell and Diebold (2002).*

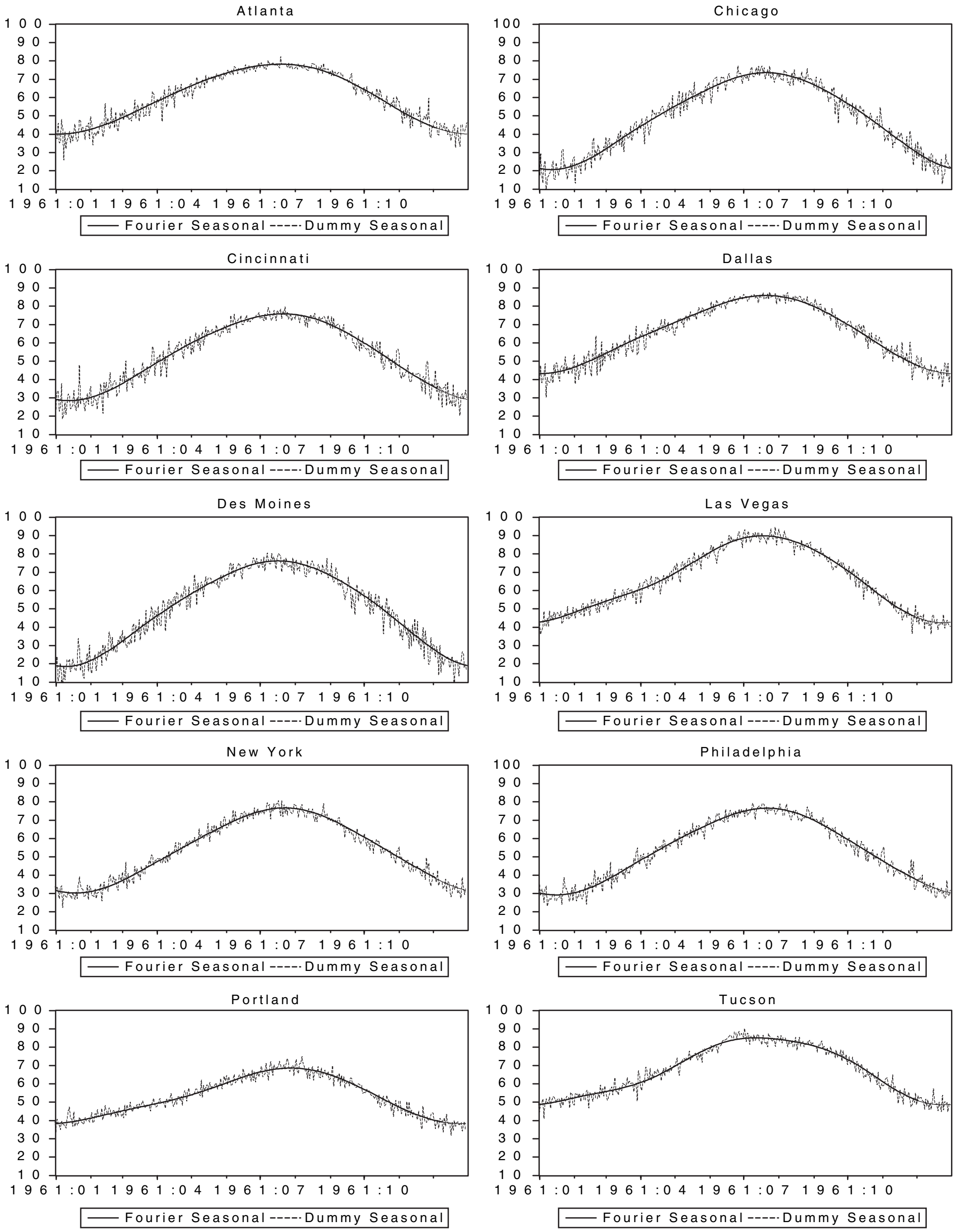

Figure 6.11 Estimated seasonal patterns of daily average temperature (Fourier series versus daily dummies). *Source: Campbell and Diebold (2002).*

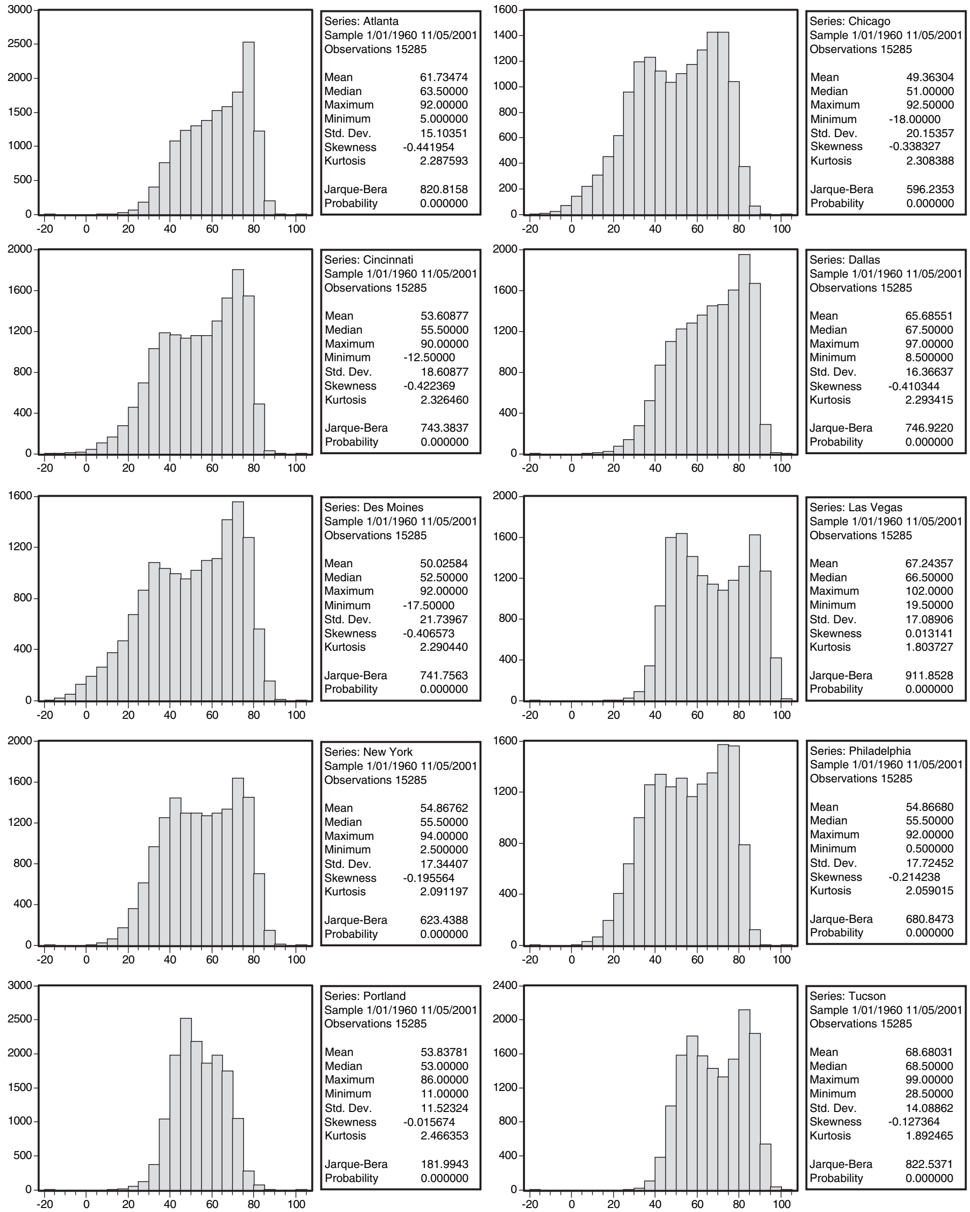

Figure 6.12 Histograms of the estimated unconditional distributions of daily average temperature in various U.S. cities from 1996-2001. *Source: Campbell and Diebold (2002).*

6.10 Pricing Weather Options in C++

A weather option such as HDD or CDD can be thought of as an Asian option because the payoff depends on the average temperature over the period. We can price these options using Monte Carlo. We can adapt the code of an arithmetic Asian option to price a weather option—e.g., an HDD—by using the average formula in (6.1) and taking the individual payoff in (6.2) and the associated summed payoffs in (6.46). For each day in the period (of M days), we need to simulate the temperature N times (the number of times we sample the temperature during the day) and compute the high and low and take the average. This procedure is simulated K times, and the average payoff is taken over all paths and discounted to generate the expected value. We simulate the weather process using an Ornstein-Uhlenbeck process.

MCPricer_calcMCAsianPriceWeather.cpp

```
vector<double> MCPricer::calcMCAAsianPriceWeather(double price,
  double strike, double DDstrike, double vol, double rate,
   double div, double T, char type, long K, long M, long N)
{
        // initialize variables
        double A = 0.0;                 // arithmetic average
        double mu = 0.0;                // drift
        int i, j, k;                    // counters
        double deviate;                 // normal deviate
        double stddev = 0.0;            // standard deviation
        double stderror = 0.0;          // standard error
        double W = 0.0;                 // weather option price
        double sum = 0.0;               // sum payoffs
        double sum1 = 0.0;
        double sum2 = 0.0;              // sum squared payoffs
        double payoff = 0.0;            // payoff
        double val = 0.0;               // option value
        double dt = (double) T/N;       // compute step size
        double a = 0.10;                // mean reversion
        double Wbar = 65;               // long-run temperature
        double minW, maxW = 0;          // min and max temperature
        double tick_size = 100;         // tick size
        vector<double> value;           // store price, std dev., etc.
        vector<double> Wvec;            // store temperatures
    mrng.sgenrand(unsigned(time(0)));   // initializer RNG

        // number of simulations
        for (k = 0; k < K; k++)
        {
                payoff = 0;
                // for each day
                for (i = 0; i < M; i++)
                {
                W = Wbar;
                        Wvec.clear();
                        Wvec.empty();
                        // number of time steps (hours) in each day
                for (j = 0; j < N; j++)
                {
```

```
                            deviate = mrng.genrand();
                                mu = -a*(W - Wbar);
                            W   = W + mu*dt + vol*sqrt(dt)*deviate;
                        // Ornstein-Uhlenbeck process
                                Wvec.push_back(W);
                }
                        // sort temperatures
                        sort(Wvec.begin(),Wvec.end());
                        minW = Wvec[0];
                        maxW = Wvec[Wvec.size()-1];
                A = 0.5*(maxW + minW);
                if (type == 'C')  // cooling days
                      payoff += tick_size*max(A - strike, 0);
                else
                      payoff += tick_size*max(strike - A,0);
                }
                sum += payoff;
                sum2 += payoff*payoff;
        }
        sum = (double) sum/(K*M);  // average over all pays and days
        sum2 = sum*sum;
        val = exp(-rate*T)*(sum);
        value.push_back(val);

        stddev = sqrt((sum2 - sum*sum/M)*exp(-2*rate*T)/(K-1));
        value.push_back(stddev);

        stderror = stddev/sqrt(K);
        value.push_back(stderror);

        value.push_back(payoff);

        if (type == 'C')
                value.push_back(tick_size*(DDstrike - payoff));
        else
                value.push_back(tick_size*(payoff - DDstrike));

    return value;
}
```

The main method is

MCPricer_weather_main.cpp

```
#include "MCPricer.h"

void main()
{
    MCPricer mcp;
    long N = 5;
    long M = 100000;

    // weather options
    double W = 50;
    double strikeW = 65;
    double volW = 0.2;
```

```
    long numDays = 30;
    long K = 10000;
    long numSteps = 100;
    double HDDstrike = numDays*strikeW;
    double mat = (double) numDays/360;
    vector<double> val;
    cout.precision(8);

    val = mcp.calcMCAAsianPriceWeather(W,strikeW,HDDstrike,volW,rate,div,
       mat,'C',K,numDays,numSteps);
    std::cout << "Weather option price = " << val[0] << endl;
    std::cout << "Std deviation    = " << val[1] << endl;
    std::cout << "Std Error =        " << val[2] << endl;
    std::cout << "Actual HDD Price = "<< val[3] << endl;
    std::cout << "HDD Payoff at Maturity = " << val[4] << endl << endl;
}
```

The is the price of an HDD with a mean reversion of 0.1, a volatility of 0.2, a long-run mean of 65°F, with $K = 10000$ simulations, $M = 30$ days, and $N = 100$ time steps.

```
Weather option price = 14.462588
Std deviation = 0.14220213
Std Error = 0.0014220213
Actual HDD Price = 433.15165
HDD Payoff at Maturity = 151684.83
```

ENDNOTES

1. Platen and West (2004), 2.
2. Alaton, P., Djehiche, B., and Sillberger, D. (2000), 1.
3. Cao, Li, and Wei (2004), 1.
4. Benth, F. and Salyte-Benth, J. (2005), 9.
5. Id., 9.
6. Alaton, P., Djehiche, B., and Sillberger, D. (2000), 9.
7. Id., 9.
8. Dornier and Queruel (2002).
9. Alaton, P., Djehiche, B., and Sillberger, D. (2000), 10.
10. Basawa and Prasaka Rao (1980), 212–213.
11. Brockwell and Davis (1990).
12. Bibby and Sorenson (1995).
13. Alaton, P., Djehiche, B., and Sillberger, D. (2000), 16.
14. Platen and West (2004), 23.
15. Alaton, P., Djehiche, B., and Sillberger, D (2000), 17.
16. Zeng, L. (2000), 77.
17. These are reported in the six columns next to the first column. The last three columns show the option value estimates under different sample lengths.
18. Cao, Li, and Wei (2004), 14.
19. Id., 14.
20. Id., 14.

21. Cao, Li, and Wei (2004), 15.
22. Campbell, S. and Diebold, X. (2000), 5.
23. Id., 5.
24. Campbell, S. and Diebold, F. (2000), 6.
25. Id., 6.
26. Id., 6.
27. Id., 6.
28. Campbell, S. and Diebold, F. (2002), 9.
29. Engle (1982).
30. Campbell, S. and Diebold, F. (2002), 9.
31. Id.
32. Id., 10.