8

Weather Derivative Modelling and Valuation: A Statistical Perspective

Anders Brix, Stephen Jewson and Christine Ziehmann

Risk Management Solutions

Weather derivatives are usually priced by analysing the historical outcomes of the underlying weather index. In this chapter we review statistical and actuarial methods for such analyses and discuss the relevance of arbitrage pricing. We look at reasons for trends in historical data and describe how to estimate and remove them. Statistical methods for modelling and validating models for weather indexes and daily temperatures are discussed, and in particular we show that traditional ARMA time-series models are not adequate for modelling daily temperatures. We show how dependencies between different indexes and different locations can be modelled and we review some of the methods for risk loading of actuarial prices.

Introduction

Weather derivatives are different from most other derivatives in that the underlying weather cannot be traded. Furthermore, the weather derivatives market is relatively illiquid. This means that weather derivatives cannot be cost-efficiently replicated with other weather derivatives, ie, for most locations the bid–ask spread is too large to make it economical to hedge a position. One of the consequences of this is that valuation of weather derivatives is closer to insurance pricing than to derivatives pricing (arbitrage pricing). For this reason it is important to base valuation on reliable historical data, and to be able to model the underlying indexes accurately.

In the future, the weather derivatives market may become more liquid, and at that point it may be possible to use other weather derivatives for hedging and thereby derive prices from the market for at least some contracts. However, the main purpose of this chapter is to review how weather derivatives are priced using historical data, and to highlight some of the challenges that arise when doing so. The presentation is statistical in its focus on the choice of models and model validation.

Before discussing the topics of detrending (removing trends from a time-series), index and daily temperature modelling, portfolio modelling and risk loading (the risk premuim added to the expected payoff in order to compensate the risk bearer for taking on risk), this chapter begins with with a discussion of the relationship between index and payoff distributions, which will be useful in the later sections.

Throughout the chapter the methods described will be illustrated with a relatively commonly traded contract: a New York LaGuardia, May 1–September 30 cooling degree-day (CDD) call option. Because more than 90% of the weather derivatives currently traded are based on temperature (WRMA, 2002), the main focus of this chapter will be on models for such indexes. Most of the index-based methods described would, however, apply to indexes based on other weather variables. The models for daily temperatures, on the other hand, would probably not apply to any other index types. The formulae are illustrated with some commonly used distributions in Panel 1.

Deriving the payoff distribution

The simplest way to understand the distribution of possible payoffs related to a weather index is to model the distribution of the index rather than the payoff. The reason for this is that limits on the payoff result in payoff distributions that are mixtures of discrete and continuous distributions in the sense that they will have discontinuities at zero and at the limit (see the following sub-section "A call option example"). Index distributions on the other hand are usually either discrete or continuous. This means that it is often possible to find a standard statistical distribution which models the index in a satisfactory way. The question of finding such a distribution is discussed in more detail in later. Alternatively, we could model daily temperatures, derive the corresponding index distribution. This will be discussed further in in the "Modelling daily temperatures" section.

Once we have an estimate of the index distribution we can derive the payoff distribution of the weather derivative. Exact and approximate expressions can be derived in many cases, but it is often simpler and faster to simulate realisations from the estimated index distribution, and convert each simulated index into a payoff. Nevertheless, knowing how the payoff distribution is derived theoretically from the index distribution can be useful for model validation and understanding which aspects of the index distribution are the most important.

In the following, I denotes the underlying index and we denote that the distribution of I has cumulative distribution function (CDF) F. As an example we will consider a call option on a degree-day index, but other contract types are treated in similar ways.

A CALL OPTION EXAMPLE

We consider a call option with strike S (in degree days), limit L (in degree-days) and tick d (in US dollars per degree-day). The payoff P of the call option is given by

$$P = d(\min(I, L) - \min(I, S)) \tag{1}$$

$$= \begin{cases} 0 & I \le S \\ d(I - S) & S < I \le L \\ d(L - S) & I > L \end{cases}$$

Using the tick we can convert the limit, L, into a USdollar limit $L_\$ = d(L–S)$, and the CDF G of the payoff can be expressed as:

$$G(P) = \begin{cases} F(S) & P = 0 \\ F\left(S + \dfrac{P}{d}\right) & 0 < P \le L_\$ \\ 1 & P > L_\$ \end{cases} \tag{2}$$

Because of the strike and limit, the payoff P is partly discrete with point masses at zero and the limit $L_\$$. Between these two points the distribution is discrete or continuous depending on whether the underlying index distribution is discrete or continuous.

The mean of the payoff can be conveniently calculated using the limited expected value (LEV) L_I for the index I[1]:

$$L_I(m) = E\, min(I, m)$$
$$= \int_{-\infty}^{m} I \, \mathrm{d}F(I) + m\,(1 - F(m))$$

Here, m is the arguement of L_I and it is seen that if we let m tend to infinity, $L_I(m)$ tends to the expected value of I. Taking the expectation of Equation (1) and using the definition of $L_I(m)$, we see that the mean payoff is given by:

$$\mathrm{E}P = d\,(L_I(L) - L_I(S))$$

This expression shows that the LEV function is one of the fundamentally relevant properties of the index distribution when pricing weather derivative options. If other moments of the payoff distribution are of interest, they can be calculated using higher-order LEV functions such as $\mathrm{E}\ \mathrm{min}(m, I)^k$.

The formulae for calculating payoff distributions and LEV functions are illustrated with some commonly used distributions in examples 1–3 in Panel 1.

Adjusting for warming and cooling trends

One of the common requirements for accurate modelling of weather indexes is that the series of historical indexes is stationary, which roughly means that the distribution of indexes does not change over time. Obviously, stationarity cannot be assumed: climate change and urbanisation may both be reflected in data. Climate change refers to variations in the Earth's climate on large spatial scales due to either natural variation or anthropogenic changes in the composition of the atmosphere. In contrast, urbanisation, which can also greatly alter a measurement station's temperature record, is a local and regional effect.

In global temperature trend studies, correction for local effects is often made either by using rural stations alone or by using rural stations to "correct" measurements from stations in urban areas. Several papers describe such studies, and their main findings are:

1. The global average surface temperature has increased by 0.6 (+/– 0.2) degrees Celsius since the late 19th century (IPCC, 2001).
2. Trends depend on the period chosen. For the US, for example, the temperature history since 1910 can be divided into three periods: a warming period until 1940, a cooling period from 1940 to 1970, and the recent warming period from 1970 to the present (Knappenberger *et al*., 2001).
3. Trends depend on location. The recent period of warming has been almost global, but the largest increases of temperature have occurred over the mid- and high latitudes of continents in the northern hemisphere. Year-round cooling is only evident in the northwestern North Atlantic and the central North Pacific Oceans (although the North Atlantic cooling appears to have reversed recently, see for example Hansen *et al.*, 1996 and IPCC, 2001).
4. Trends for maximum and minimum temperatures are different. The diurnal range, the difference between daily maximum and daily minimum temperatures, is decreasing, although not everywhere. On average, minimum temperatures are increasing at about twice the rate of maximum temperatures (see for example Easterling *et al*., 1997 and IPCC, 2001).
5. A recent study of Knappenberger *et al.* (2001) shows that the trends are not uniform; cool days are much warmer than they used to be, whereas warm days are not.

Urbanisation effects are best demonstrated by comparing temperature time-series from neighbouring stations where one is from an area with little urban change while major changes have been made to the surroundings of the other. Such a comparison

PANEL 1

DERIVING THE PAYOFF DISTRIBUTION

Example 1

Consider the case where the underlying index follows a *log normal distribution* with parameters μ and σ2, ie, log *I* follows a normal distribution with mean μ and variance σ^2. The payoff CDF is:

$$G(P) = \begin{cases} \Phi\left(\dfrac{\log S - \mu}{\sigma}\right) & P = 0 \\ \Phi\left(\dfrac{\log (S + P / d) - \mu}{\sigma}\right) & 0 < P \leq L_\$ \\ 1 & P > L_\$ \end{cases}$$

where Φ is the standard normal CDF. The LEV function is:

$$L_I(m) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \Phi\left(\frac{\log m - \mu}{\sigma} - \sigma\right) + m(1 - G(m))$$

Hence the expected payoff is:

$$E\,P = d \exp\left(\mu + \frac{\sigma^2}{2}\right)\left(\Phi\left(\frac{\log L - \mu}{\sigma} - \sigma\right) - \Phi\left(\frac{\log S - \mu}{\sigma} - \sigma\right)\right) + dL(1 - G(L)) - dS(1 - G(S))$$

Example 2

If the index follows a *normal distribution*, with mean μ and variance σ^2, the situation is a bit more complicated because the limited expected moments are less tractable. The payoff CDF is simple:

$$G(P) = \begin{cases} \Phi\left(\dfrac{S - \mu}{\sigma}\right) & P = 0 \\ \Phi\left(\dfrac{S + P / d - \mu}{\sigma}\right) & 0 < P \leq L_\$ \\ 1 & P > L_\$ \end{cases}$$

The LEV function is given by

$$L_{\mu,\sigma}(m) = \sigma L_{0,1}\left(\frac{m - \mu}{\sigma}\right) + \mu\Phi\left(\frac{m - \mu}{\sigma}\right) + m\left(1 - \Phi\left(\frac{m - \mu}{\sigma}\right)\right)$$

Here the LEV $L_{0,1}$ for the standard normal is given by

$$L_{0,1}(m) = \int_{-\infty}^{m} \frac{u}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du$$

which can be calculated numerically.

Example 3

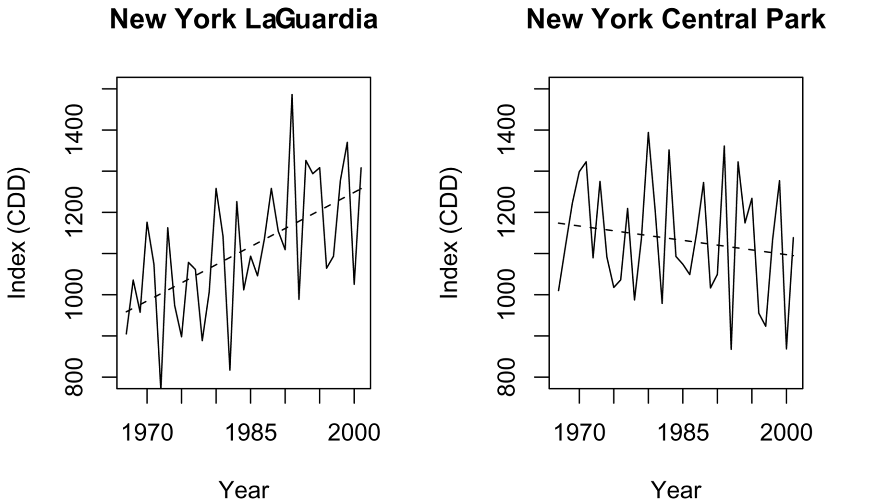
The general formulae above apply equally well when the underlying index follows a discrete distribution. Consider the case where I follows a *negative binomial distribution*. Although it is easy to calculate numerically, there is no closed-form expression for the CDF of this distribution. In the following we will denote the CDF of the negative binomial with mean $r(1-p)/p$ and variance $r(1-p)/p^2$ by $F_{r,p}$. The payoff CDF is given by Equation (2) with F replaced by $F_{r,p}$ and the LEV function is given by:

$$L_{r,p}(m) = \frac{rp}{1-p} F_{r+1,p}(m) + m(1 - F_{r,p}(m))$$

can only be made if the two stations are in the same microclimate region. As an example, this chapter will compare the CDD index for New York LaGuardia Airport with the corresponding index for Central Park, New York. These two stations are in the vicinity of each other, but whereas much has been built in the LaGuardia area over the last 30–40 years and it is at the water's edge, not much has changed structurally over the same period around Central Park, in the heart of the New York City. Figure 1 shows the historical CDD indexes for the two locations with linear trends overlaid. Visually, the difference between the two plots is striking, and t-tests of the significance of the slopes of the trendline reveal that the Central Park trend is not significant (p=34%), while the LaGuardia trend is highly significant (p=0.06%).

For the purposes of this chapter, it is not important to distinguish between local and global trends because we are interested in removing the combined trend. There are many models in the statistical literature for estimating distributions with trends. The way the trend is incorporated is often linear or multiplicative in the mean (for example, an index is typically modelled as trend plus noise for a normal distribution, and trend multiplied by noise for a log-normal distribution), and is usually chosen on the basis of mathematical convenience rather than reflecting reality (since it can be difficult to find a 'realistic' model for many applications and since the appropriateness of the model must be validated anyway). This section separates trend estimation from distribution estimation on the basis that this simplifies the calculations, and because we can then treat all distributions in the same way.

1. Historical indexes for New York LaGuardia Airport (left) and Central Park, New York (right). Linear trends have been superimposed.



Trend estimation is done in essentially the same way for daily temperatures and annually compounded indexes. Because seasonal effects can complicate the process of daily detrending, detrending of annually compounded indexes is described first, and daily detrending later. Non-stationarity due to seasonality is discussed under the heading of "Capturing seasonality of temperatures" below.

There can be other sources of non-stationarity than trends, such as, relocation of measurement stations or equipment and sudden changes to the surroundings of stations. The effects of such measurement discontinuities can be better dealt with by first enhancing the underlying data, as described in detail in Chapters 5 and 6. This chapter is concerned with the more gradual changes that remain after those that are associated with measurement discontinuities have been removed.

INDEX DETRENDING

The assumption of the trend model used here is that an index I_i can be represented as a sum of a trend R_i and a random variable e_i:

$$I_i = R_i + e_i, \qquad i=1, \ldots, n$$

e_i are assumed to be independent and identically distributed with mean zero. The detrended indexes, $\tilde{I}_i$, are then defined as

$$\tilde{I}_i = \mathrm{I_i} - \hat{\mathrm{R}} + \hat{\mathrm{R}}_\mathrm{n} \qquad (3)$$

where $\hat{\mathrm{R}}$ and $\hat{\mathrm{R}}_\mathrm{n}$ are the estimated trends for indexes $\boldsymbol{i}$ and $\boldsymbol{n}$.

In this way, the mean of all indexes are shifted to the estimated mean of the last index. Often the contract will commence a year (or more) after of the end of the historical indexes. If the trend is thought to continue after the last historical data point it can be extrapolated to year $n + k$, where k is the number of years to extrapolate forward. We then replace R_n by R_{n+k} in Equation (3), and get the k-year ahead forward detrended indexes. For example, suppose we are looking at data from a station which has experienced large growth in urbanisation in its surroundings in recent years and where the urbanisation is still continuing. If we are considering a contract for the winter, the most recent historical data would be from the previous winter and we would need to extrapolate the estimated trend in order to capture the trend introduced by continuing urbanisation.

Parametric trends

In this chapter, trends are assumed to be smooth and vary slowly in time since we have assumed that jumps due to for example station relocations have been removed. Therefore, it is often reasonable to approximate trends by parametric curves such as linear or polynomial functions. The standard way of estimating the parameters of the trend is by ordinary least squares (OLS), ie, minimising the sum:

$$\sum_{i=1}^{n} (I_i - R_i)^2$$

With y_i denoting the year of index i, the trend R_i could then be parameterised by, for example:

$$R_i = a + b\,y_i \qquad \text{(linear)}$$

$$R_i = a + b\,y_i + c\,y_i^2 \qquad \text{(quadratic)}$$

$$R_i = a\exp(b\,y_i) \qquad \text{(exponential)}$$

OLS estimates are known to be sensitive to extreme observations, so if outliers are present, a more robust estimation procedure may be needed – least absolute deviations for example (ie, minimise $\sum_{i=1}^{n} | I_i - R_i |$. See Huber (1981) for a detailed treatment of robust statistics).[2]

Non-parametric trends

Sometimes it may be desirable to use a non-parametric trend if there is reason to believe that parametric trends do not provide a satisfactory approximation for the period considered. In such situations various non-parametric methods are available (see, for example, Bowman and Azzalini, 1997). The simplest is called the "moving average" method, where the trend in year i is estimated as the average of the neighbouring years:

$$R_i = \frac{1}{2w + 1} \sum_{i=-w}^{w} I_{i+w}$$

The number of neighbouring years, $2w + 1$, is usually called the "window", and the years may be weighted such that years closer to the base year contribute more than years that are further away. The main disadvantage of moving average estimation is that it does not extrapolate the trend beyond the last historical year.

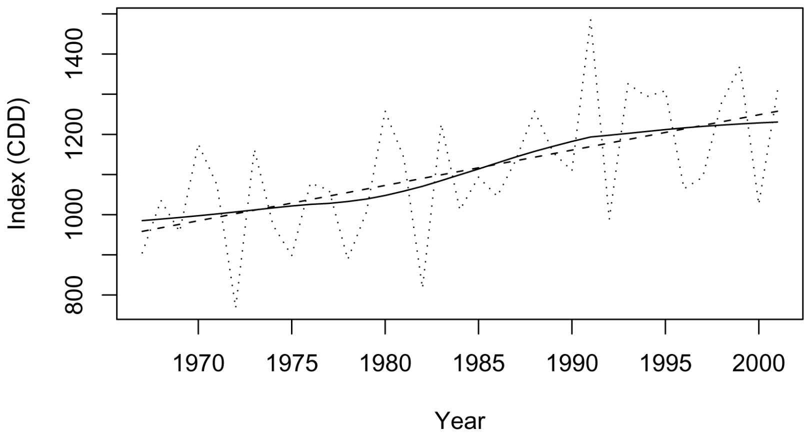
An alternative that allows extrapolation is the "loess" method (Cleveland and Devlin, 1988), which is based on local parametric regressions. Linear loess, for example, estimates the trend for year i by weighted linear regression, with most weight on nearby years. Loess is known to have better theoretical properties than moving average estimation, especially close to the edges of the observation window.

Figure 2 shows estimated linear trend and loess trends for the CDD example used in this chapter. The difference between the two trendlines ranges from –26 to 26 CDDs.

DAILY DETRENDING

Detrending of daily temperatures can be done using the methods described above. However, local or global warming effects may have different magnitudes in different seasons (Hansen *et al.*, 1996), thus create different trends at different times of the year. This is not a problem when modelling annual indexes since such seasonality does not appear, and similarly non-parametric trends adapt to each season. Parametric trends, on the other hand, may need to be adapted to vary by time of the year. One way this can be done is to estimate linear trends separately for each month

2. Estimated linear (dashed) and loess (solid) trends for the CDD example. The original index values are connected by dotted lines.



of the year and interpolate to get a trend for each day. The seasonal linear trend at day t then becomes:

$$R_t = a + b_t t$$

where b_t is the slope on day t, which will follow an annual pattern.

Even if the analysis is based on annual indexes, it may be worth using daily detrending when considering short contracts. When using index-based trends, the estimated trend for a weekly contract may be implausibly different from the estimated trend for the same contract the following week. This point is illustrated in Figure 3, which shows how estimates of linear trends of average temperature vary by week when they are estimated using only data from that week. By detrending daily values from a period longer than the week of the contract, we use information about the trend in the surrounding weeks, and thus get a more stable estimate.

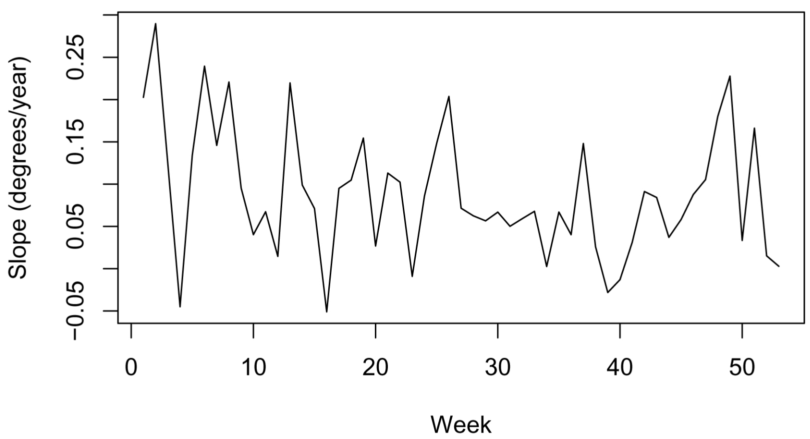
HOW MANY YEARS OF DATA SHOULD BE USED?

One of the reasons for linear and higher-order polynomial trends is that they provide good approximations to smooth trends over periods of a reasonable number of years. What is a reasonable number of years? The answer varies by location, since many trends are due to local effects, but also because of spatial variation in global warming. Based on backtest studies it has been found that, on average over many stations and many years, a reasonable number of years would be between 15 and 25 if a linear index trend is used.[3] If more data is used the quality of the linear trend line deteriorates due to non-linearity, and if less is used, there is too little data to get an accurate estimate of the index trend and distribution. It must be stressed, though, that this guideline is valid only on average, and that the number of years that is appropriate for individual stations may lie outside this range.

VALIDATING THE TREND

So far this section has discussed how to estimate a given trend, but avoided giving guidelines on how to chose a trend and how to validate it. One (and arguably the best) way to do this is by graphical checks. First, a plot of the index values can give some idea about whether a trend might be present or not. Second, after estimating the trend, a residual plot should be made. The residuals are the difference between the indexes and the trend, and should be scattered around zero without any systematic variation by time, ie, the residuals must show no clear trend, no change in spread over time and there must not be a clear tendency for the points to be mainly

3. Linear trends of average temperature at New York LaGuardia Airport for each week of the year.



negative or positive. For all the methods outlined above, the residuals should also be approximately normally distributed. With only a few index values, validation can be a difficult task, and any external information on when changes in trends could have occurred should be used when choosing a trend estimate. For example, many temperature measurements are made at airports, or in city centres, so information about when the airport or city has grown is useful for determining when a trend could have started.

Weather index modelling

Having described detrending of historical indexes in the previous section, we will now discuss which types of statistical distributions are appropriate for modelling the detrended indexes. Both parametric and non-parametric methods will be considered and a discussion on model validation will follow.

NON-PARAMETRIC INDEX MODELLING

The simplest form of non-parametric distribution is the empirical (historical) distribution of the indexes. In the actuarial literature use of this distribution is known as "burn" analysis. However, given that we usually deal with relatively few historical indexes, the empirical distribution can become very rugged and sometimes have several modes (peaks). If the distribution of the index is thought to be smooth or unimodal, then this may not be realistic. Instead the empirical distribution can be smoothed by a process called "kernel smoothing", whereby the probability density function f of the index distribution taken at the point x is estimated by the expression

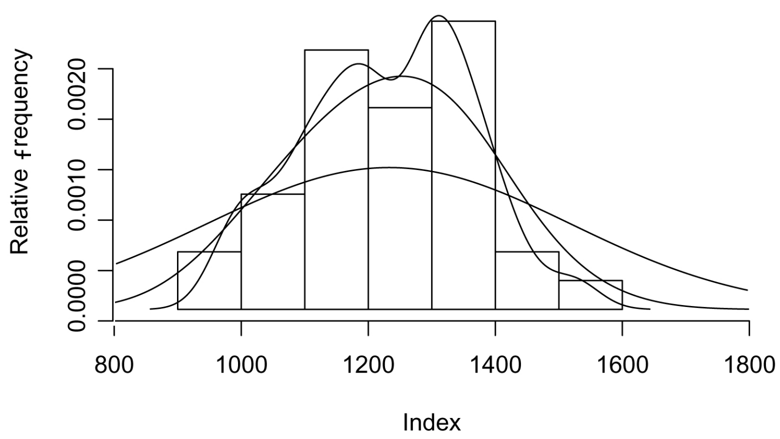
$$\hat{f}(x) = \frac{1}{n} \sum_{i=-1}^{n} \frac{1}{h} \, k\left(\frac{x - \tilde{I}_i}{h} \right)$$

Here, k is a probability density function (PDF), and the degree of smoothing is determined by the bandwidth, h. The effect of kernel smoothing is illustrated in Figure 4, where three Gaussian kernel density estimates have been superimposed on a histogram. Whereas the choice of smoothing function k is not very critical, the bandwidth selection is extremely important for the overall shape of the estimated distribution: the larger h, the more smoothing is obtained.

PARAMETRIC INDEX MODELLING

Even with a large degree of smoothing the kernel distribution may put too much weight on the historical data; for example, we may not believe in multimodality, the

4. Histogram with Gaussian kernel-smoothed densities overlaid for three different bandwidths (40, 100 and 250).



historical distribution may not look smooth enough or we may want to extrapolate further beyond the historical observations than is possible by kernel smoothing. If so, we can use parametric distributions such as normal, gamma or Poisson, depending on what type of index is considered. A normal distribution is often an appropriate choice for degree-day and average indexes, especially for longer contract periods. The reason for this is the central limit theorem, which states that, under very general conditions, the sum of a number of outcomes (like daily degree-days or daily temperatures) approximately follows a normal distribution (see, for example, Casella and Berger, 2002). Similarly, it can be shown that extreme events, such as the number of days with temperature exceeding a high threshold, will be approximately Poisson-distributed if they are close to independent (see for example, Coles, 2001). However, extreme weather events tend to occur in clusters, and so are often better modelled by a negative binomial distribution (see for example, McCullagh and Nelder, 1989).

In general, parameters of the distribution are most efficiently estimated by the maximum likelihood method, ie, the parameter estimates are chosen to maximise the PDF (probability mass function for discrete distributions) as a function of the parameters. Alternatively, parameter estimates can be obtained by deriving expressions for the moments, and matching these with empirical estimates. For the normal distribution the two methods are equivalent and amount to calculating the empirical mean and variance.

VALIDATING INDEX DISTRIBUTIONS

The sparsity of data with which one is often faced in weather index modelling can make it difficult to distinguish between a bad fit due to sampling error and a bad fit due to wrong choice of model. This makes it particularly important to estimate how much variation can be expected due to sampling error and to apply a variety of model checks. The following paragraphs describe a number of graphical techniques, show how these can be made more quantitative through simulation and discuss goodness-of-fit tests.

PP-, QQ- and LL-plots

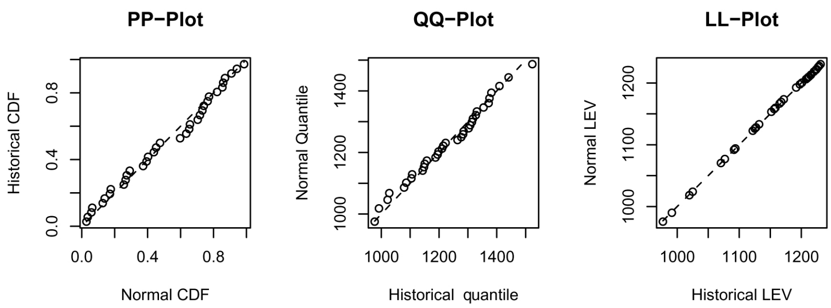
Two traditional ways of validating the fit of a distribution are by comparing the model PDF and CDF with the histogram and the empirical CDF, respectively. Such comparisons are, however, not without difficulties. For histograms the number of bins and the bin width have to be chosen, and for CDFs it can be difficult to distinguish between the different S-shapes that these usually take. Instead, it is customary in statistics to look at so-called PP- and QQ-plots, which contain the same information as the CDF but allow easier comparison of distributions.

A PP-plot is a plot of the empirical CDF against the model CDF; if the model is good, then the points should be close to a straight line with slope of one and intercept of zero. This way, the problem of comparing S-shapes is turned into a comparison of straight lines, which is much easier. Note that a PP-plot is also useful for getting a qualitative evaluation of lack of fit. If the points fall on a straight line with an intercept different from zero this indicates that the mean is wrong. If the points fall on a straight line with slope different from one this indicates wrong variance. The left panel of Figure 5 shows a PP-plot for the CDD example for New York.

Because all CDFs start at zero and end at one, points in the tails of the distributions on the PP-plot will tend to be close to a straight line. This makes it difficult to evaluate the quality of the fit to the tails of the distribution. A QQ-plot, where model quantiles are plotted against empirical quantiles, is a way around this problem. Again the points should lie close to the straight line with slope of one and intercept of zero if the fit is good. The advantage of QQ-plots is that a bad fit in the tails of the distribution will show more clearly. The middle panel of Figure 5 shows a QQ-plot for the CDD example.

While PP- and QQ-plots provide good tools for evaluation of the overall fit, it is

5. Validation plots for a normal distribution for the CDD example



often useful to check the fit of specific characteristics of the distribution that are more closely related to the actual purpose of the modelling. Since we are interested in modelling payoffs, it is important to capture accurately those features of the index distribution which affect the basic characteristics of the payoff distribution. Hence, if we are modelling options or swaps, we should compare the model and empirical LEV functions. As before, we prefer to look for straight lines when validating the fit, so we plot the two against each other. The plot created this way is called an LL-plot, and is shown in right panel of Figure 5 for the CDD example.

Envelopes

The validation plots described in the previous section are purely descriptive, and it can sometimes be hard to see when deviations from a straight line are due simply to sampling variation. In order to make them more quantitative we can add confidence intervals to the plots. It is usually not possible to derive exact expressions for such confidence intervals – we use simulation envelopes instead. This is done by simulating samples of the same length as the historical data from the estimated model; if the model is good, then the CDF of the samples should fall around the historical CDF.

Simulation envelopes for the PP-plot are obtained as follows:

Step 1. Simulate a sample of the same length n as the historical data from the estimated model.
Step 2. Calculate the empirical CDF from the sample at each historical index.
Step 3. Repeat Steps 1 and 2 K times and store the results.
Step 4. Sort the calculated CDF at each value of the historical indexes.

To achieve 90% confidence intervals we could simulate, say, K=100 samples and pick out the fifth lowest and the fifth highest CDF values for each historical index.[4] The envelopes created this way are pointwise confidence intervals; there is 90% probability that the historical CDF at a given point will fall within the envelopes. Because points on the simulated CDFs are highly dependent this does not mean that the probability of the full historical CDF falling within the envelopes is 90% to the power of n.

Similarly for the QQ-plot and the LL-plot, we calculate the quantiles corresponding to the quantiles of the historical indexes, and LEV at each historical index. (Figure 10 shows an example of a QQ-plot within envelopes.)

GOODNESS-OF-FIT TESTS

Apart from graphical checks several goodness-of-fit tests are available, of which the most common are chi-square (χ^2), Shapiro–Wilks, Kolmogorov–Smirnov and Anderson–Darling. Because of the small number of historical indexes that are usually available, none of these tests are very powerful, ie, they will rarely reject a bad model. Graphical checks, such as the ones described above, often give a much better idea of how appropriate the model is.

- ❑ The chi-square test can be used for checking the validity of any distribution. The test is done by grouping the observations into intervals and comparing the expected number of observations in each interval with the observed number. Because the test relies on grouping the observations, it requires a substantial number of observations in order to give a powerful result (a general rule of thumb for when to use this test is that the expected number of observations in each group must be at least five).
- ❑ The Shapiro–Wilks test can be used only for normal distributions (and log-normal by transforming the observations with the logarithm). The test is quite powerful and gives a good indication of whether a normality assumption is reasonable.
- ❑ Kolmogorov–Smirnov is a classical test which compares the empirical CDF with the model CDF using the maximal vertical difference between the two. The test is not very powerful, but should it result in a low test probability then there is good reason not to rely on the model.
- ❑ Anderson–Darling can be considered as a modification of the Kolmogorov–Smirnov test which compares the CDFs over the whole range of the distribution. In contrast to the Kolmogorov–Smirnov test, the Anderson–Darling test probability is model-specific and hence the test is more powerful. Both this test and the Kolmogorov–Smirnov test can be used only for continuous distributions.

Modelling daily temperatures

The most common approach to analysing weather derivative contracts is to fit a distribution to the detrended annual indexes. However, often we have only a few historical values from which to estimate this distribution, resulting in significant estimation uncertainty. This uncertainty could be reduced if the data were used more efficiently, and the index approach has some inefficiencies, as shown by the following examples:

- ❑ For a US cooling degree-day (CDD) index, the index approach uses only information about how far above 65°F the temperature is. It does thus not distinguish between days where the temperature is far below 65°, and days where the temperature is just below 65°.
- ❑ Event indexes only use data from days on which events occurred. Data on all other days is discarded.
- ❑ One-week contracts only use data for that week of the year. Data from other weeks is discarded.
- ❑ For some indexes, in particular indexes relating to short periods and extreme events, it may not be possible to find a suitable model for the index distribution.

These problems could be alleviated if the underlying daily temperature distribution could be modelled. Furthermore, a temperature modelling approach would make it easier to include forecasts in the weather derivative pricing process (see Chapter 10).

To see how much more efficiently data can be used in a daily model compared with an index model, consider the May–September CDD contract for New York LaGuardia Airport. We estimate the index distribution from a daily temperature model and an index model using the same amount of historical data. Both estimates of the index distribution will have an error due to estimation uncertainty, and these errors can be quantified by looking at confidence intervals (envelopes) around the estimated CDFs. Figure 6 shows the estimated CDFs with 90%-confidence envelopes using a normal index distribution and a daily temperature model, respectively.

The daily model used for Figure 6 is the CJB model to be described later, but it must be noted that almost any model for daily temperatures would show this apparent decrease in estimation uncertainty. This is because a daily model uses the data more efficiently than an index model and because the comparison is based on the assumption that the model is correct (the simulation envelopes are created using the estimated model for both the index and the daily model). However, a bad daily model could also produce significant bias, and hence reduce the value of increased estimation accuracy. For this reason it is extremely important to validate a daily model before using it for weather derivatives pricing; validation methods are discussed in detail later.

Another reason for a thorough validation of daily validation models is that a daily model may give a different index distribution from that given by the historical index values (and from a normal distribution). Since a daily model is using data more efficiently than an index model we can place more weight on the daily model results, but only if it has been thoroughly validated.

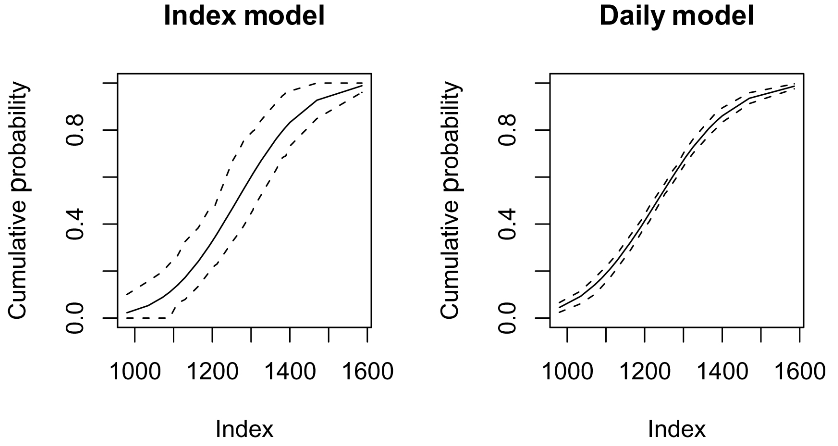
Statistical modelling of daily temperatures has been a research subject for decades (see for example the review in Wilks and Wilby, 1999) and, more recently, a number of papers on weather derivative pricing that propose models for temperature have been published (Alaton *et al*. (2002), Brody *et al.* (2002), Caballero *et al*. (2002), Cao and Wei (2000), Davis (2001), Diebold and Campbell (2001), Dischel (1998), Moreno (2000) and Torró *et al*. (2001)). Common to all of the above referenced papers except Brody *et al.* (2002) and Cabellero *et al.* (2002) is that they use ARMA-type models which will be discussed in more detail in the following subsections.

We will start by discussing the basic steps of building a model for daily temperatures and then show how statistical validation techniques are applied. The model validation will show that ARMA-type models fail to validate well on two points:

1. The modelled auto-correlation function decays too quickly relative to reality.
2. The residuals deviate markedly from their theoretical distribution.

For this reason we also show validation results for the model discussed in Caballero *et al*. (2002) which, in general, validates much better than ARMA models for daily temperatures. We conclude the section with a brief discussion of some of the problems that still remain to be solved and which, to the authors' knowledge, apply to all published daily temperature models.

6. Potential gain in accuracy from daily modelling over index modelling. Normal distribution index CDF (left) and index CDF derived from CJB model (right), both with 90%-envelopes (dotted).



CAPTURING SEASONALITY OF TEMPERATURES

One basic principle underlying statistical models for stochastic processes is that the data should be stationary, ie, the distribution should be invariant over time. Apart from trends, daily temperatures have some obvious non-stationarities, namely those due to seasonal variation.

Seasonal variation undoubtedly affects daily temperature in many complicated ways (eg, in all the moments), and it will probably never be possible to remove all types of seasonality, let alone to check that it has been done. In the following it is assumed that most seasonality is due to seasonal variation in the mean and the standard deviation. In mathematical terms we assume that the time-series of daily temperatures can be decomposed as follows

$$T_i = m_i + s_i T'_i \tag{4}$$

where T_i is the temperature, m_i is the mean temperature, s_i is the standard deviation and T'_i is a mean zero variance one variable that is known in meteorology as the "anomaly" for day i. The anomalies T'_i are assumed to be a stationary time-series, for which a large number of candidate models exist.

Like estimation of trends, seasonality estimation can be either parametric or non-parametric. The following section discusses a number of such methods for removing the seasonality in the mean. The seasonal variance can be estimated and removed by exactly the same method applied to the square of the difference between the temperature and the seasonal mean. For this reason we only describe the methods for estimating the seasonal mean.

Non-parametric approaches

The simplest way to estimate the seasonal mean is to average each day of the year over the historical data period, ie, the estimate for January 1 is the average of January 1 in all years and so on. When using this method, special care must be taken of leap days. If the seasonal mean estimated this way is considered too ragged it can be smoothed, by means of kernel smoothing, for example. One problem is that even with smoothing, results are often too jumpy.

Parametric approaches

Another way of estimating the seasonal mean is by parameterising the mean temperature using trigonometric functions. This has the advantage that leap days are easy to take account of. We can estimate the m_i using ordinary regression with m_i given by

$$m_i = \alpha \cos\left(\frac{2\pi}{365.25} i + \omega\right) \tag{5}$$

where α is the amplitude and ω is the phase of the seasonal mean.

Alternatively we can estimate m_i in the frequency domain. In both cases more harmonics can be included in the parameterisation of m_i if necessary (ie, using cosines to produce

$$m_i = \sum_{k=1}^{K} \alpha_k \cos\left(\frac{2k\pi}{365.25} i + \omega\right)$$

While the parametric form in Equation (5) is very simple, it can be justified by plots of power spectra of temperature anomalies which show surprisingly clear peaks at 365.25 days (and its harmonics).

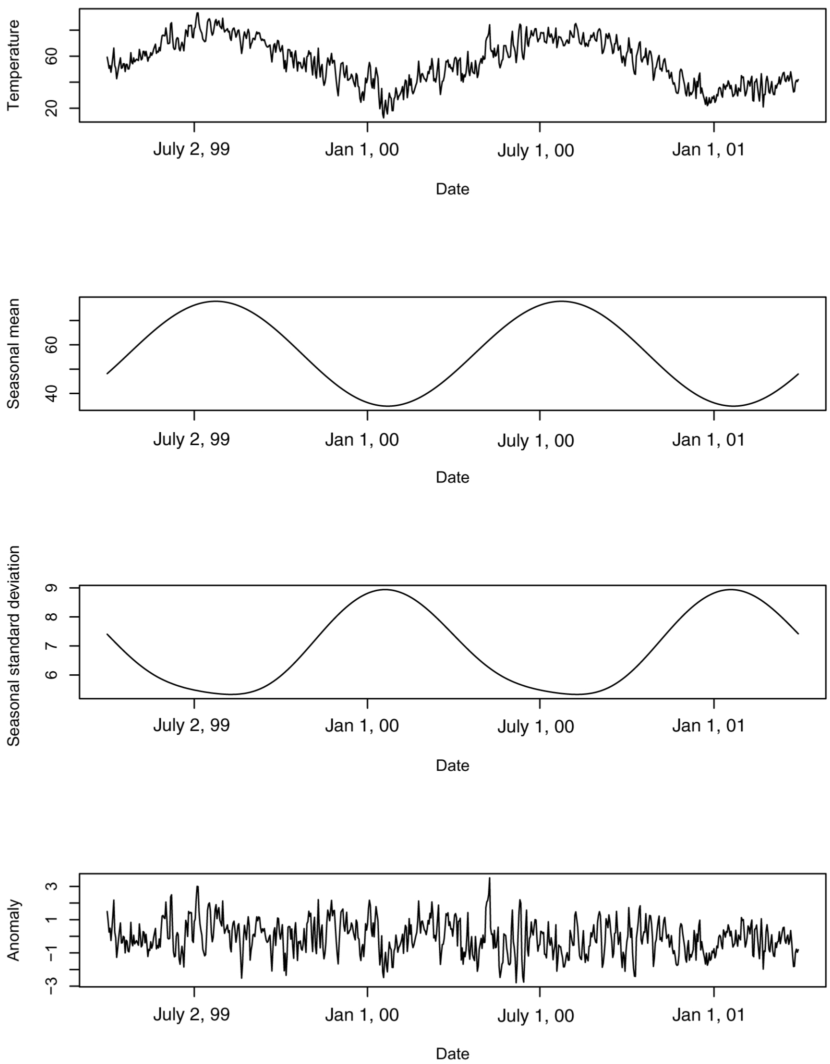
The result of applying a seasonal linear trend and parametric seasonal cycles for mean and standard deviation with two harmonics to the New York LaGuardia Airport example can be seen in Figure 7, which shows how the detrended daily temperatures are decomposed into seasonal mean, seasonal standard deviation and anomalies.

ANOMALY MODELLING

The biggest challenge in modelling daily temperatures is to find an appropriate model for the anomalies. We have seen that the index distribution's LEV function plays an important role in deriving the moments of the payoff distribution. Similarly there are functions of daily temperatures which are important for capturing moments of the index distribution from a model for daily indexes. For example, it is important to model both the seasonal cycle and the autocorrelation of temperatures accurately in order to get a good estimate of the index mean and variance of degree-day or average-temperature indexes. Consider, say, a CDD index. Using the representation in Equation (4), and assuming that the daily temperature never falls below 65°F, we have the following expressions for the mean and variance of the index:

$$\mathrm{E}\ I = \sum_{i=1}^{n} m_i - n65$$

7. Detrended daily average temperature, seasonal mean, seasonal standard deviation and anomalies.



$$V\ I = \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j \gamma(i-j) \tag{6}$$

Here, n is the number of days in the contract and γ is the autocorrelation function (ACF) of the anomalies. From Equation (6) it can be seen that we would underestimate the index variance if the model ACF, γ, decays too quickly to zero.

What kind of ACF behaviour do daily temperatures show? Dependencies in day to day temperature arise from complicated atmospheric, oceanic and land surface processes, many of which evolve slowly. In particular, ocean circulation can have cycles from years to decades and even centuries. It would thus not be surprising if temperature were to show dependencies on long timescales. As we shall see later, the ACF of temperature anomalies does indeed show relatively slow decay.

NON-PARAMETRIC TIME-SERIES MODELLING

Non-parametric time-series modelling can be done by resampling from the observed anomalies. The resampled time-series can then be used to estimate distribution statistics of interest, weather derivative indexes, for example. Resampling should not be done completely at random, since we would then lose all dependencies in the time-series. Instead we could resample blocks of anomalies and concatenate these. If the blocks are sufficiently long and if we have enough data, this should result in simulated time-series which would mimic the behaviour of the original time-series. The length of the blocks should be chosen so that the statistics of interest are not too affected by the "breaks" between blocks. For example, block lengths could be chosen to equal the contract length in the case of weather derivatives. However, it may be difficult to capture the distribution of a time-series by resampling, even if long blocks are used. See Davison and Hinkley (1997) for a survey of non-parametric resampling methods.

PARAMETRIC TIME-SERIES MODELS

Traditional time-series models are parametric models, and the simplest example is probably a first order autoregressive (AR(1)) process. For such a process the temperature at one day is given by a linear dependency on the temperature the previous day with some random noise added:

$$T'_i = \beta T'_{i-1} + \varepsilon_i$$

Here T'_i denotes the anomaly at day i and the ε_i are independent and identically distributed random variables following a mean zero normal distribution. Despite its simplicity an AR(1) process can be a useful model for a variety of problems, but is unfortunately too simple for daily temperature anomalies (one reason being that the ACF decays too quickly). A simple extension of the AR(1) process provides us with a flexible class of models, known as ARMA models, which can be used to approximate any stationary time-series model.[5] The definition of an ARMA(p, q) process is:

$$T'_i = \phi_1 T'_{i-1} + \ldots + \phi_p T'_{i-p} + \theta_1 \varepsilon_{i-1} + \ldots + \theta_{i-q} \varepsilon_{i-q} + \varepsilon_i$$

The interpretation of the model is that the temperature today depends in a linear way on the temperatures on the previous p days through the parameters $\phi_1, \ldots, \phi_p$. Just as in the AR(1) case random pertubations are added to reflect the fact that we do not expect the temperature today to be a perfect linear function of the past p days' temperatures. The $\theta_1, \ldots, \theta_q$ are parameters used to express linear dependence between the random pertubations.

Although ARMA models can, in theory, approximate any time-series model, it is actually not the best class of models for temperature anomalies. This is partly because of the slow decay of the ACF of temperature anomalies (see Figure 8 and the subsection on validation using ACFs), which by Equation (6) implies that an ARMA model would result in underestimated variances for degree-day indexes – a fact that has been recognised by many participants in the weather market for some time.[6] If we were to model a slower decay of the ACF using an ARMA model, we would need far more parameters than would be feasible to estimate in practice. A more parsimonious choice of model (in the following referred to as the CJB model) that preserves much of the ARMA model flexibility but has slower decaying ACF is described in Caballero *et al.* (2002).

In the following sections we illustrate classical statistical methods for model validation, which provide more evidence on how ARMA models fail to provide adequate modelling of daily temperatures and how the CJB model overcomes many (but not all) of the problems associated with ARMA models.

Validation of daily temperature models

Modelling daily temperatures is a difficult task and careful validation of models must be undertaken before putting them into practical use for pricing weather derivatives. In the following subsections we show some simple tools for evaluation of daily temperature models for weather derivative pricing. The methods are illustrated on the New York example, using an ARMA(3,1) model and the CJB model.

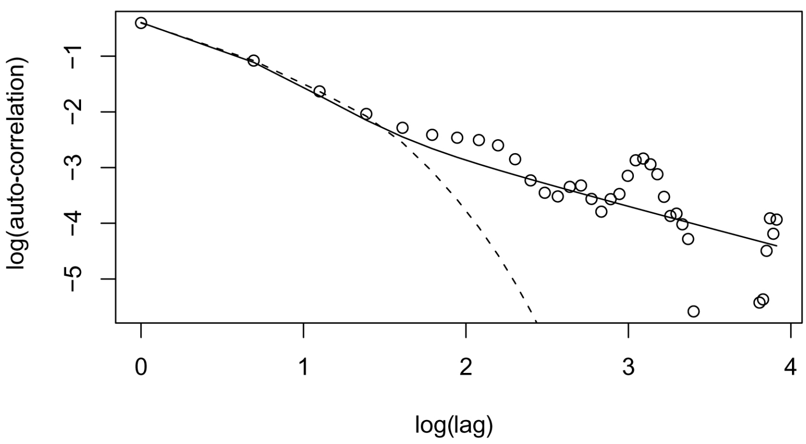
AUTOCORRELATION FUNCTION

Equation (6) and the discussion following it highlighted the importance of capturing the ACF of the temperature anomalies, making the comparison of the model ACF with the empirical ACF an appropriate step in validation of the model. However, since all ACFs start at one and usually decay quickly over the first few lags, comparison can be difficult using the usual representation of an ACF. Instead, we compare ACFs by plotting the logarithm of the ACF against logarithm of the lag. The plot in Figure 8 is a plot of this type and emphasises what we have already mentioned: that the ARMA model ACF decays too fast to capture the behaviour of the empirical ACF.

DISTRIBUTION OF RESIDUALS

Residuals are the difference between the observed anomalies and the predicted anomalies based on the past observations, ie, the historical one-step prediction errors. The parameters used for the prediction are estimated from the full time-series. For most models, including the two considered here, it is possible to derive theoretical expressions for the distribution of the residuals. We can thus check the

8. Comparison of ACFs on a log–log scale: empirical ACF (circles), ARMA(3,1) (dashed) and CJB model (unbroken).



time-series model by comparing the residuals from the model with their theoretical distribution. For non-parametric methods the theoretical distribution of the residuals is unknown and hence cannot be checked.

The left plot of Figure 9 shows a residual QQ-plot for the ARMA(3,1) model, and the right plot shows a QQ-plot for the CJB model. We see that whereas the CJB model shows consistency with the historical anomaly distribution, the ARMA(3,1) model results in a residual distribution with far too light tails.

DISTRIBUTION OF ANOMALIES

The next step in validation is the verification of the distribution of the anomalies. For mathematical convenience most time-series models assume a normal distribution, so a natural first step is to compare the empirical anomaly distribution with that of a normal through PP- and QQ-plots. If a normal distribution is not suitable, it is often possible to find a transformation that makes the anomaly distribution approximately normal.

Models also exist for time-series with non-normal behaviour such as heavy-tailed or skewed distributions. For such models, however, it can be difficult to verify when the model is stationary and determine exactly what the marginal distributions are.

ACCURATE MODELLING OF WEATHER DERIVATIVE INDEXES

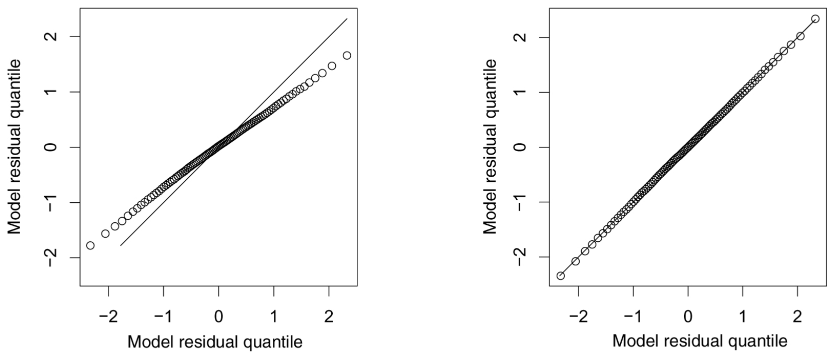
While it is important that the time-series of temperatures is modelled accurately, the ultimate goal is to model weather derivative indexes. The transformation of temperatures to a degree-day or an event-specific index (such as a critical-day index) is non-linear, and hence it is not clear how small model deficiencies at the temperature level will propagate to errors in the index distribution. This, however, is not a problem for a period-average temperature index. We need to validate the index distribution, derived from a daily model, using the same tools as we used for validating the index model. As an example, the two plots in Figure 10 show the QQ-plots for the index distribution resulting from using an ARMA(3,1) and the CJB model.

Imposing 90%-confidence envelopes on the plots we see that the CJB model falls within the envelopes, but that the ARMA model is questionable because of the relatively large number of points that fall outside or on the boundary of the envelopes.

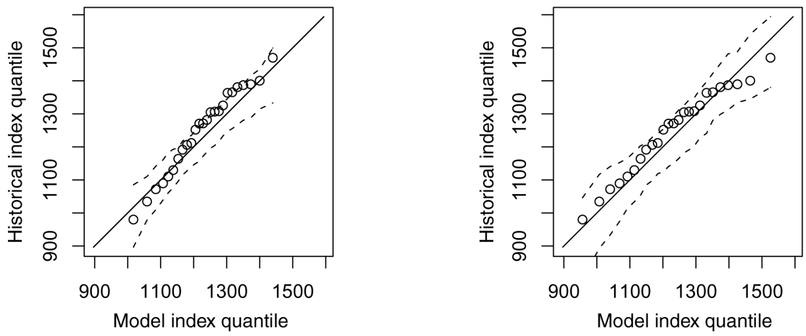
Some outstanding issues in daily temperature modelling

We have seen that ARMA models are not adequate for daily temperature modelling and that the CJB model is a better alternative. However, there are still some

9. QQ-plot for the residuals using an ARMA(3,1) model (left) and CJB model (right) for New York LaGuardia Airport.



10. QQ-plot for the index distribution for New York LaGuardia Airport using an ARMA(3,1) model (left) and CJB model (right).



unresolved issues which means that even the CJB model may not be adequate in all situations. In particular some locations show strong seasonal variation beyond that of mean and standard deviation, which in turn means that both ARMA and CJB models validate badly for these locations.

Another issue is that modelling daily temperatures for several locations simultaneously is complicated by the fact that the spatial correlation structure is non-stationary. This makes it difficult to exploit all the information in daily temperatures and as a result dependencies between locations are usually modelled on an index level instead (see the following section).

Capturing index dependencies

Weather indexes often show strong geographical dependencies – a fact that must be captured by our modelling approach. One reason for this is that pricing of a weather contract is generally done in the context of the current portfolio: the more risky the portfolio gets by adding the contract, the less that contract is worth. Furthermore, portfolio modelling is essential to valuing a book of contracts and the principles for risk loading described in the following section rely crucially on appropriate joint modelling of the contracts in the portfolio. To model dependencies, various tools exist – mostly based on correlations. However, correlations are not just correlations!

The usual notion of correlation is linear correlation, which is what is calculated in spreadsheets and other standard software packages. Linear correlation is only an appropriate measure of dependency if the joint distribution of the indexes is close to Gaussian or elliptical (Embrechts *et al.*, 2000). If this is not the case, linear correlations may fail to accurately capture the dependency between indexes, especially in the tails of the distribution. Moreover, the correlation coefficient need not have a maximum of one and a minimum of minus one in the non-Gaussian case. Instead the minimum and maximum linear correlation will depend on the distribution and its parameters, which makes it more difficult to evaluate the degree of dependence. Fortunately other methods for capturing dependence exist, two of which are discussed here: rank correlation and copulas.

RANK CORRELATION

Rank correlation is a alternative way of measuring dependencies between distributions. All values between one and minus one are possible whatever the distribution, which makes interpretation of rank correlation easier than linear correlation in the non-Gaussian case.

Rank correlation is calculated by transforming all indexes by their empircal CDF and then computing the linear correlation between the transformed indexes. This is equivalent to computing the linear correlation between the rank of the sorted indexes (hence the name) and the method can be seen as a non-parametric approach to capturing dependence between indexes.

COPULAS

The notion of copulas is based on the same fundamental idea as the rank correlations, namely that dependency between distributions is compared after the distributions have been transformed to uniforms: a copula is simply the CDF of a multivariate distribution with uniform marginals. For example the Gaussian copula is the CDF given by transforming each variable of a multivariate Gaussian distribution by its CDF. Rank correlations can be thought of as a special type of copula, where the dependence is given by the empirical correlation of the observed indexes transformed to uniforms. In general, however, a copula is given as a parametric CDF, and the parameters of the copula must either be estimated from data or guessed from intuition. Guessing is usually not a good idea, but due to lack of data it may be necessary. Likewise the preference of one specific copula over another can be hard to justify from data, and is often made on the basis of mathematical convenience.

Risk loading principles

Every risk has its price, and the difference between this and a fair price is called "risk loading". How is an appropriate risk loading calculated? The principles outlined below are general principles for how to obtain a price of risk based on risk measures such as standard deviation and quantiles. Most of the principles can also be adapted to other risk measures.

The price of a contract does not always depend directly on how risky it is. If, for example, we know that we can buy a certain contract from somebody else in the market at a given price it would be risk-free to offer the contract at a higher price. This way of determining a price is driven by the market's perception of risk rather than one's own. This is discussed further in the paragraphs that follow.

SIMPLE ADDITIVE RISK LOADING

The simplest risk loaded price, P_r, of a contract is the expected outcome plus λ times the risk, where λ is a risk factor defining the speculator's appetite for risk. Ignoring discounting, the pricing formula is:

$$P_r = \mathrm{E}P + \lambda R \qquad (7)$$

where R is the risk measure and $\mathrm{E}P$ is the expected payoff. A measure of risk is needed, and it is customary to use standard deviation or variance of the payoff for this purpose. Alternatively we could use the difference between two quantiles such as the median and the 5% quantile, whereby we get a risk measure which is more akin to Value-at-Risk.[7]

Equation (7) also allows the pricing of the contract against the full portfolio. To do this we change the risk measure R to be a measure for the additional risk that is taken on by trading the contract. Changes in any of the risk measures above could be used for this purpose. For example, we could use the difference between the standard deviation of the portfolio payoff with and without the contract as risk measure R.

INVESTMENT EQUIVALENT PRICING

Most companies will have ways of allocating capital to contracts based on their own internal measures of risk. Had the capital not been allocated to a contract it could have been invested in other assets. This gives the speculator another way of pricing a contract: the return on the allocated capital should be at least what would be expected on an investment in a similar risk elsewhere. This is the idea behind so-called Risk Adjusted Return On Capital (RAROC) (Nakada *et al.*, 1999) and investment equivalent reinsurance pricing (Kreps, 1999).

Consider the case where an option is sold at a price Pr. The price consists of the discounted expected payoff and risk loading, L:

$$Pr = \mathrm{E}P/(1 + r_f) + L \quad (8)$$

where r_f is the risk-free interest rate. In addition to the premium income we allocate an amount of capital A to the contract and invest $Pr + A$ in risk-free securities. The idea is now that the average return on allocated capital should equal the return from an investment in an alternative instrument with the same risk. The expected Return on Allocated Capital (RAC) is thus defined by

$$(1 + RAC)\, A = (1 + r_f)\,(Pr + A) - \mathrm{E}P$$

Combining this with Equation (8) we get the following expression for the loading

$$L = (\mathrm{RAC} - r_f)/(1 + r_f)A$$

Using the estimated payoff distribution we can get an idea of how risky the option is and use this to find a reasonable target RAC which, in turn, gives the risk loading. Several variations of this theme are possible, see Kreps (1999).

Other topics

We have seen in this chapter that valuation of weather derivatives is a broad subject where many factors must be taken into account. While we have tried to cover the majority of topics there are still some that we have left out and many which could have been further elaborated.

ARBITRAGE PRICING

Arbitrage pricing is the standard way of pricing financial derivatives in a liquid market. As noted in this chapter's Introduction most weather derivative contracts are not yet liquid enough to justify arbitrage pricing but some are now traded several times a day. This makes it worthwhile to consider how arbitrage pricing for weather derivatives may be done.

The basic principle of arbitrage pricing is that the cost of a derivative is the cost of creating and managing a portfolio which replicates the payoff of the derivative contract at maturity. The active management of the replicating portfolio is what is usually called dynamic hedging. However, the underlying index of a weather derivative cannot be traded and hence weather derivatives cannot be hedged this way. Instead we could replicate the payoff of a weather derivative using other weather derivatives. While this is possible in principle, the illiquidity of the current weather market makes it prohibitively expensive for most types of contracts. However, should the market become liquid enough, several arbitrage strategies would be possible.

One strategy would be to use properties of forecasts to derive a theory for how the market would price a weather swap. Since swaps can be used to hedge options, this gives us a way of obtaining an arbitrage price for an option. The ideas behind this are discussed in detail in Stephen Jewson's contribution to the End Piece.

Some authors advocate hedging of weather derivatives with other derivatives such as power or gas derivatives (Geman, 1999). However, such hedges are not likely to be complete hedges and we are thus left with a basis risk. In order to find the price of the basis risk we must model the weather derivative and the underlying of the alternative hedge jointly, for example using methods similar to those previously discussed in this chapter. Note also that whereas HDDs are highly correlated with gas consumption, they are usually very little correlated with gas prices, which is one of the reasons for trading weather derivatives in the first place.

OTHER WEATHER VARIABLES

Another challenging topic is modelling of daily weather variables other than temperature, such as of wind speed or precipitation. These relatively uncommon underlyings (compared with temperature) are dealt with thoroughly in Chapter 3.

EXTREME VALUE THEORY

One subject that has not been mentioned is the use of extreme value theory (EVT) for evaluating risk and estimation of the distribution of extreme indexes. While this is an interesting point of discussion, the current state of the market is such that weather derivatives are mostly written on non-extreme risk, and as such EVT is less applicable.

Conclusion

This chapter has described how weather derivatives can be valued by statistical and actuarial methods based on historical data. It has given an overview of all the important topics: estimation and adjustment of trends, modelling and validation of detrended historical weather indexes and daily temperatures, accounting for index dependencies and risk loading of expected payoffs.

As trading liquidity increases for certain contracts, aspects of arbitrage pricing will become more important and in some cases replace actuarial methods. However, for most contracts, the actuarial methods described in this chapter will continue to be the main, and only reasonable, valuation approach that can be used.

1 *The LEV function is often used in insurance for calculating mean loss to an excess of loss reinsurance layer.*
2 *By "robust" we mean methods which, at the cost of accuracy, are less affected by outliers than traditional methods.*
3 *Based on the authors' own study using 50 years of data from 200 US stations.*
4 *In statistical literature the observed sample is often included in the K simulations since the hypothesis modelled is a sample from the correct model. In practice this distinction does not make a lot of difference.*
5 *It can be shown that for any autocovariance function* $\gamma(\cdot)$ *such that* $h \to \infty \Upsilon(h) = 0$, *and any integer* $k > 0$ *it is possible to find an ARMA process with autocovariance function* $(\cdot)$ *such that* $(h) = \gamma(h)$, $h = 0, 1, \ldots, k$.
6 *This is the authors' own experience based on discussions with the major market participants over the last three years.*
7 *Value-at-Risk (VAR) is a quantile (typically 5%) in the modelled profit and loss distribution over a specified time period, and is one of the most common measures of risk in finance.*

BIBLIOGRAPHY

Alaton, P., B. Djehiche and D. Stillberger, 2002, "On Modelling and Pricing Weather Derivatives", *Applied Mathematical Finance*, Forthcoming.

Bowman, A, and A. Azzalini, 1997, *Applied Smoothing Techniques for Data Analysis*, (Oxford University Press).

Brockwell, P., and R. Davis, 1991, *Time Series: Theory and Methods*, (New York: Springer-Verlag).

Brody, D. C., J. Syroka and M. Zervos, 2002, "Dynamical Pricing of Weather Derivatives", *Quantitative Finance*, Forthcoming.

Caballero, R., S. Jewson and A. Brix, 2002, "Long Memory in Surface Air Temperature: Detection, Modelling and Applications to Weather Derivative Valation", *Climate Research*, Forthcoming.

Cao, M., and J. Wei,. 2000, "Equilibrium Pricing of Weather Derivatives", Research Paper, Deptartment of Economics, Queen's University, available at: http://qed.econ.queensu.ca/pub/faculty/cao/papers.html.

Casella, G., and R. L. Berger, (2002), *Statistical Inference*, (Pacific Grove, CA: Brooks/Coles Publishing).

Cleveland, W., and S. Devlin, 1988, "Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting", *Journal of the American Statistical Association*, 83, pp. 596–610.

Coles, S., 2001, *An Introduction to Statistical Modelling of Extreme Values*, (New York: Springer Verlag).

Davis, M., 2001, "Pricing Weather Derivatives by Marginal Value", *Quantitative Finance*, 1, pp. 1–4.

Davison, A., and D. Hinkley, 1997, *Bootstrap Methods and Their Application*, (Cambridge University Press).

Diebold, F. X., and S. D. Campbell, 2001, "Weather Forecasting for Weather Derivatives", Working Paper, Department of Economics, University of Pennsylvania, available at: http://www.econ.upenn.edu/Centers/pier/Archive/01-031.pdf.

Dischel, R., 1998, "Black-Scholes Won't Do", Risk Magazine and Energy and Power Risk Management Report on Weather Risk, http://www.financewise.com/public/edit/energy/weatherrisk/wthr-options.htm.

Easterling, D., B. Horton, P. Jones, T. Peterson, T. Karl, D. Parker, M. Salinger, V. Razuvayev, N. Plummer, P. Jamason and C. Folland, 1997, "Maximum and Minimum Temperature Trends for the Globe", *Science*, 277, pp. 364–7.

Embrechts, P., A. McNeil and D. Straumann, 2000, Correlation and Dependency in Risk Management: Properties and Pitfalls, in: M. A. H. Dempster and H. K. Moffatt (eds) *Risk Management: Value at Risk and Beyond*, (Cambridge University Press).

Geman, H., 1999, "The Bermudan Triangle: Weather, Electricity and Insurance Derivatives, in: H.Geman (ed), *Insurance and Weather Derivatives*, (London: Risk Books).

Hansen, J., R. Ruedy, M. Sato and R. Reynolds, 1996, "Global Surface Air Temperature in 1995: Return to Pre-Pinatubo Level", *Geophysical Research Letters*, 23, pp. 1665–8.

Huber, P. J., 1981, *Robust Statistics*, (New York: John Wiley & Sons).

IPCC, 2001, "Climate Change 2001-The Scientific Basis", Technical Report, Technical Summary of the Working Group I Report.

Jewson, S., and A. Brix, 2001, "Sunny Outlook for Weather Investors?", *Environmental Finance*, February, pp. 28-9.

Knappenberger, P., P. Michaels and R. Davis, 2001, "Nature of Observed Temperature Changes across the United States during the 20th Century", *Climate Research*, 17, pp. 45–53.

Kreps, R., 1999, "Investment Equivalent Reinsurance Pricing", in: *Proceedings of the Casualty Actuarial Society*, Volume LXXXVI, (Arlington, VA: CAS).

McCullagh, P., and J. A. Nelder, 1989, *Generalized Linear Models*, (London: Chapman and Hall).

Moreno, M., 2000, "Riding the Temp", Research Paper, Speedwell Weather Derivatives, Available at: http://www.weatherderivs.com/Papers/Riding.pdf.

Nakada, P., H. Shah, H. Koyluoglu and O. Collignon, 1999, "P&C RAROC: A Catalyst for Improved Capital Management in the Property and Casualty Insurance Industry", *Journal of Risk*, 1, pp. 52–70.

Torró, H., V. Meneu and E. Valor, 2001, "Single Factor Stochastic Models with Seasonality Applied to Underlying Weather Derivatives Variables", Research Paper, Facultat d'Economia, Universitat de Valencia, available at: http://www.efmaefm.org/htorro.pdf.

Wilks, D.S., and R. L. Wilby, 1999, "The Weather Generation Game: A Review of Stochastic Weather Models", *Progress in Physical Geography*, 23, pp. 329–57.

WRMA, 2002, "The Weather Risk Management Industry: Survey Findings for November 1997 to March 2001", Report prepared for Weather Risk Management Association by PriceWaterhouseCoopers, Available at http://www.wrma.org.