Meteorol. Appl. 00, 1–13 (2003) DOI:10.1017/S1350482703001099

The use of weather forecasts in the pricing of weather derivatives

Stephen Jewson[1] & Rodrigo Caballero[2]
[1]*Risk Management Solutions, 10 Eastcheap, London, EC3M 1AJ, UK*
Email: x@stephenjewson.com
[2]*Department of the Geophysical Sciences, University of Chicago, 5734 S. Ellis Ave., Chicago, IL 60637, USA*
Email: rca@geosci.uchicago.edu

We discuss how weather forecasts can be used in the pricing of weather derivatives and derive results for the most important types of weather index and contract. We show that calculating the expected payoff of linear contracts on linear indices requires only forecasts of the mean temperature over the contract period. Calculating the expected payoff of linear contracts on non-linear indices requires forecasts of both the mean and the distribution of temperatures, but not of the dependence between temperature distributions on different days. Calculating the expected payoff of non-linear contracts requires forecasts of the full multivariate distribution of temperature over the whole contract. For contracts that extend beyond the end of available forecasts, correlations between the forecast and post-forecast periods must be taken into account when estimating this distribution. We present two methods by which this can be achieved, both of which combine information from climatological models of daily temperature with information from probabilistic forecasts.

1. Introduction

There is little doubt that weather and climate forecasts are of value to society, but there is growing awareness that the full potential economic value of forecasts has yet to be realised (Palmer 2002; Freebairn & Zillman 2002). One way to enhance the economic value of forecasts is to incorporate them into risk assessment tools that allow economic actors, be they public or private, to directly assess the financial loss or gain to them from future weather or climate events. Here, we explore the possibilities for doing this in a specific context, that of weather derivatives. These are financial contracts that allow companies (or other entities) to insure themselves against the adverse effects of fluctuations in weather (Zeng 2000). The classic example is that of a gas supply company, which will lose money in warm winters. The company may enter into a weather derivative contract with, say, a bank. The contract stipulates that, at a certain date, one party will pay the other an amount which is entirely determined by measurements of the weather. For instance, the contract might be based on the mean temperature over the period from November to March. If the mean temperature over this period turns out to be lower than normal, then the gas company will make money in its gas business but will lose money on the weather contract, and vice versa if the temperature is higher than normal. In this fashion, the company smooths its revenues from year to year, creating welcome financial stability. Clearly, weather derivatives play a very similar role to traditional insurance; the main advantages over insurance are that future income can be insured, and that the company need not go through a lengthy claim procedure each time it incurs a loss.

Many weather derivatives are traded long before the start of the contract and long before there are any useful forecasts which can indicate the likely weather during the contract period. For instance, contracts for the winter period may be traded in the preceding spring and early summer. In this case, only historical observational data are required for derivative valuation. We will refer to values of contracts calculated at this stage as the *par* values. In order to calculate these values various decisions must be made such as the number of years of historical data to use, what trends to remove, whether to extrapolate trends, and what statistical models to fit to the data.

It is also common for weather derivatives to be priced just before and during the period of the contract. There are two main reasons for this. The first is that weather derivatives are traded at these times. This can be for economic hedging reasons, or purely for speculation. The second is that companies that have traded a weather derivative often need to track the value of the derivative as the weather during the contract period progresses: this is known as 'marking' the contract. It is for the pricing of weather derivatives just before the start of, and during, the contract period that forecasts can be of value and this article describes the methods by which forecasts can be incorporated into the pricing process. We concentrate on medium-range weather forecasts, assuming either that no seasonal forecasts are available or that they are incorporated into the pricing at some stage not discussed here. The reasons for making such a separation of weather and seasonal forecasts are pragmatic: these forecasts are presented in different ways, and are available from different sources. For instance, weather forecasts are usually available as daily values, while seasonal forecasts are given as monthly or seasonal values. This forces the users of forecasts to design separate methods by which these forecasts are incorporated into weather pricing models. We will, however, briefly describe a unifying framework which allows all types of forecasts to be included in the same prediction of future temperatures.

In Section 2 we define some common types of weather derivatives, while in Section 3 we review the kinds of weather forecasts that can be used in the pricing of weather derivatives, and compare single, ensemble and probabilistic forecasts. In Section 3.3 we discuss the kinds of statistical models that are used for calculating the par values of the fair swap strikes, fair option prices, and the distributions of swap and option outcomes. We then discuss the weather forecast-based pricing of three types of weather derivatives contract, moving from simple to complex contract structures. First, in Section 4, we consider linear swap contracts based on cumulative temperature. In Section 5 we consider the same contract but based on heating degree days. Finally, in Section 6, we consider the general case of a non-linear contract on any index which motivates the methodologies presented in Section 7. These are the main new results of this paper. Section 8 presents a summary of our results and conclusions.

2. Types of weather derivatives

Weather derivatives can be based on any meteorological index, but cumulative heating and cooling degree days, cumulative temperatures and average temperatures are particularly common. We will focus the discussion in this article on cumulative heating degree days and cumulative temperatures. The methodologies we will describe for heating degree days can equally well be applied to cooling degree days, and those for cumulative temperatures can equally well be applied to average temperatures.

The number of heating degree days on a single day is defined (in Europe) as

$$y_i = \max(18 - T_i, 0), \qquad (1)$$

where T_i is the daily average temperature in °C. A cumulative heating degree day index x over a period of M consecutive days is defined as the sum of the heating degree days on each day during that period:

$$x = \sum_{i=1}^{M} y_i. \qquad (2)$$

Similarly, a cumulative temperature index over M consecutive days is defined as

$$x = \sum_{i=1}^{M} T_i. \qquad (3)$$

The rule which determines how much is paid to whom as a result of the final outcome of the weather index can take any form, but most contracts are either *swaps* or *vanilla options*. In a swap contract there is no premium paid at the start of the contract: the two sides simply enter into an agreement. At the end of the contract money is paid in either direction according to whether the index is above or below a predetermined value known as the *strike*, and the amount paid is proportional to the distance from the strike. Sometimes, the amount paid is not allowed to exceed a fixed upper limit, in which case the swap is said to be 'capped'. For an uncapped swap, the amount of money changing hands is a linear function of the index, while for capped swaps it is non-linear. The upper panels of Figure 1 show the money that changes hands in uncapped and capped swap contracts as a function of the final index, from the point of view of the buyer of the contract (who is defined as the party that receives money if the index ends up with

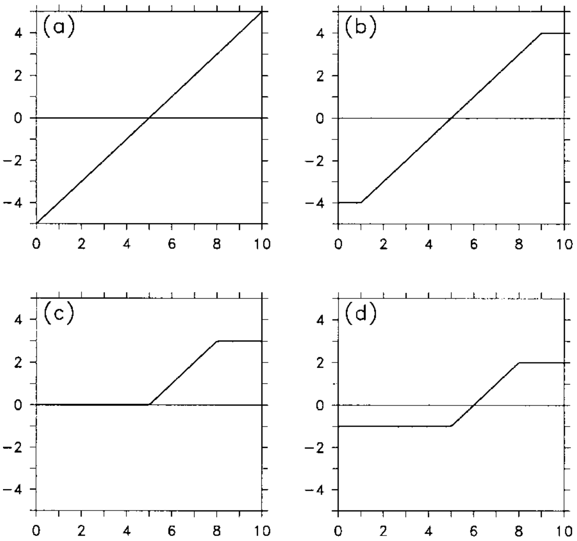


Figure 1. *Some examples of payoff functions for swap and option contracts. The horizontal axis shows the final realised value of the weather index on which the contract is based, and the vertical axis shows the amount that would be paid by the seller to the buyer as a function of that index. (a) An uncapped swap contract, with strike at 5 and a tick (the slope of the line) of 1. (b) A capped contract, the same as (a) but with caps at ±4. (c) A call option with strike at 6 and a cap at 3. (d) The same call option, but inclusive of the premium paid by the buyer, which in this case was taken to be 1.*

a high value, and pays money if it ends up with a low value).

A representative type of vanilla option is a *call* option. A call option is similar to an insurance contract in that the buyer pays a premium at the start of the contract, and may receive a payout at the end. It provides insurance against high values of the index. For values of the index above the *strike*, the seller pays the buyer an amount proportional to the difference between the index and the strike, with the constant of proportionality given by the *tick*. As with swaps, the payout is often capped. The lower left panel of Figure 1 shows the money that changes hands at the end of a call option as a function of the final index, from the point of view of the buyer of the contract. The lower right panel of Figure 1 shows the net exchange of money including the premium.

The simplest approach to the valuation of swap contracts consists of estimating the value of the strike at which the expectation of the distribution of financial outcomes of the contract is zero, known as the 'fair strike' (Note that this definition of fair is nothing more than a reasonable convention: another reasonable convention would be that 'fair' is defined by the strike which makes the *median* of the distribution of payoffs zero.). Traded strikes are usually close to this fair value but can move away from it due to changes in supply and demand. Valuation of option contracts usually starts with estimating the expected payout of the contract, known as the 'fair premium' or 'fair price'. Since the buyer of the contract is usually purchasing a service and the seller is providing a service by taking on the risk of having to pay out, options are usually bought for a premium that is significantly above the fair value. In addition to calculating fair strikes and fair prices, it is also important for the buyers and sellers of weather derivatives to understand the entire distribution of possible outcomes of contracts that they trade. They may manage their risk by monitoring, for example, the standard deviation, or the 5% quantile, of this distribution.

3. Weather forecasts

We will consider weather forecasts that predict temperatures for today, and for 11 days into the future (so there are 12 days in each forecast). Since weather derivative indices are based on measurements made at individual weather stations, such forecasts must be downscaled to the correct space-time location to be of use. Such downscaled forecasts are routinely available on a commercial basis and will be the starting point of our discussion. The two main kinds of forecast we will consider are single and ensemble forecasts. We will see that both are useful, and we will discuss how both can be used to derive *probabilistic* weather forecasts, that is, forecasts which give probabilities for every possible outcome of weather over the next 11 days. In order to keep things simple and tractable, and for the sake of illustration, we will generally assume that temperature is normally distributed. This is a good assumption in some cases, but not for all (see Jewson & Caballero 2002 for examples both ways). With some extra work, the methods we will describe below can be extended to other distributions.

A probabilistic forecast for daily values of temperature for today and the next 11 days consists of the specification of a single multivariate distribution. If the temperatures are normally distributed, then we are dealing with a multivariate normal, which is fully specified by the means of the marginal distributions together with the 12×12 covariance matrix (see, for example, Reiss & Thomas 1997: 164).

3.1. Single forecasts

A single forecast consists of a single time series of 12 temperatures. There are two main forms this forecast can take, which differ in terms of the variance in time and hence in how they can be interpreted statistically. First, a single forecast could be a possible track for the future weather, which we will call a *realistic track* forecast. As such it would have the same unconditional variance as the actual temperatures for that location and time of year, i.e. the variance in time of the forecast would be the same as the variance in time of actual temperatures. The second interpretation is that the time series could represent the best estimate of the mean of the future range temperatures, which we will call a *conditional mean* forecast. In this case the forecast would not necessarily be a realistic track for future weather, and is likely to decay towards climatological temperature for that time of year as the forecast progresses. Such forecasts are defined by their variance, which is set so as to minimise root mean square forecast error rather than to equal the variance of actual temperatures. Conditional mean forecasts are also known as a 'tempered' forecasts (Leith 1974). The mathematics of these two kinds of forecasts, and their relationship to ensembles, is discussed in detail in Jewson & Ziehmann (2002).

Given a set of past realistic track forecasts, one can easily derive a conditional mean forecast by building a linear least squares minimising regression model between the past forecasts and reality (von Storch & Zwiers 1999; Jewson et al. 2002b). Realistic track forecasts are then converted to conditional mean forecasts by passing them through this regression model. If it is not known whether a particular forecast is a realistic track or a conditional mean, such a regression can be used to test the forecast. A conditional mean forecast will have regression coefficients close to one. Realistic track forecasts are usually produced from suitably downscaled, single integrations of a numerical weather prediction (NWP) model. Conditional mean forecasts

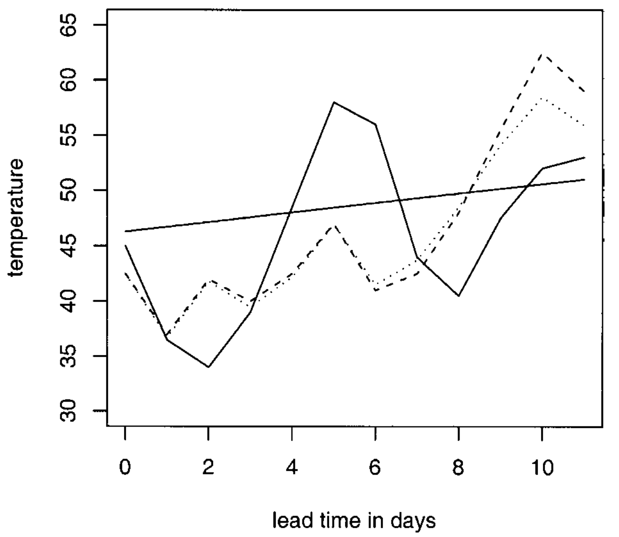


Figure 2. *A single track forecast for Chicago made on 10 April 2000 (dashed line), along with the actual temperature (solid line, °F), the climatological mean (straight solid line), and a conditional mean forecast calculated from the single track forecast (dotted line).* Q1

Table 1. *The regression coefficients between forecast and reality, calculated using three months of single track forecasts. Note that these coefficients are less than one beyond five days, implying that the original forecast is not a conditional mean forecast. These are the regression coefficients that were used to calculate the conditional mean forecast in Figure 2.*

Lead time	Damping coefficient
0	1.01
1	1.03
2	1.04
3	1.07
4	1.06
5	1.03
6	0.93
7	0.82
8	0.72
9	0.74
10	0.66
11	0.61

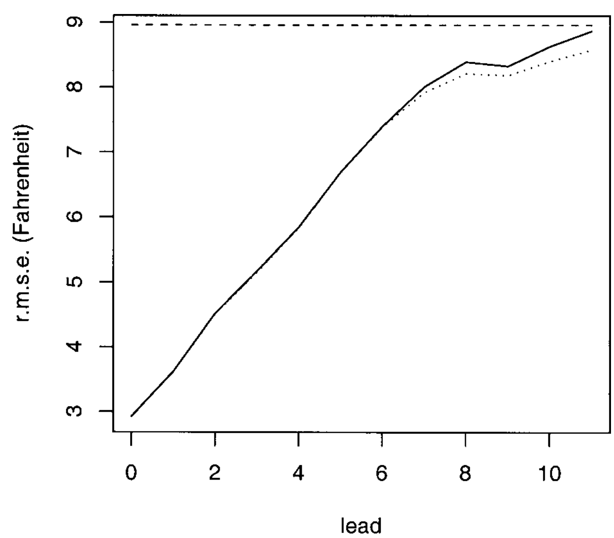
are produced from realistic track forecasts using the regression model described above, or using a regression model applied to the mean of an ensemble forecast (see below). For weather derivative pricing, our main interest in single forecasts is in conditional means rather than realistic tracks.

Conditional mean forecasts can be used to derive probabilistic forecasts. By taking a history of past forecast values, we can derive error statistics that allow us to put a distribution of possible outcomes around future forecasts. If we assume normality of the forecasts the error statistics fix the variances and covariances of the multivariate forecast distribution. Since forecast error statistics can vary from season to season this should be taken into account in such a model. One possible shortcoming of this simple approach for creating probabilistic forecasts is that it does not allow the variances or covariances in the forecast to vary with the state of the atmosphere.

As an example of how a single forecast can be used to make a probabilistic forecast, Figure 2 shows a single forecast (obtained from a commercial forecast provider) for Chicago made on 10 April 2000, along with actual temperature and the climatological mean (based on 30 years of linearly detrended data) for the forecast period. Using the previous three months of forecast data, we build a regression model between the forecast and reality. The coefficients are shown in Table 1. We can see that, beyond day 5, this forecast is not a correctly damped conditional mean forecast (the coefficients are not close to one). In order to correct this, we pass the forecast for 10 April 2000 through the regression model, giving a new forecast which is closer to climatology for days 6–10. This forecast is also shown in Figure 2. The root mean square (RMS) errors for both forecasts estimated over the three months for April–June are given in Figure 3. We can use the RMS errors for the corrected forecast to put a distribution on the conditional mean: the 5% and 95% quantiles of this distribution are shown in Figure 4, along with the conditional mean and the actual temperatures.

3.2. Ensemble forecasts

Ensemble weather forecasts for daily temperature consist of a set of realistic track forecasts covering the same 12-day span. Taken together, they give an indication of the range of possible future outcomes of the weather. If one assumes that the individual tracks



Q1

Figure 3. *RMS error for a single track forecast (solid line), and for a conditional mean forecast (dotted line) derived from it. The RMS error for the conditional mean forecast, is, by construction, lower than that of the original forecast.*

Q1

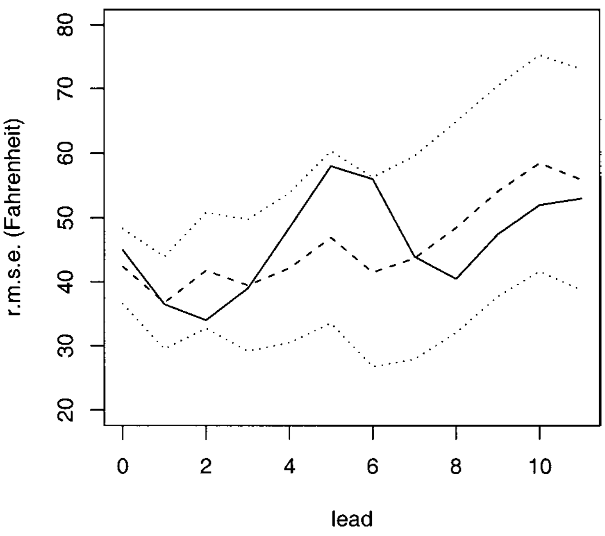


Figure 4. *A probabilistic forecast derived from a conditional mean forecast using past forecast errors. The dashed line shows the conditional mean forecast, the solid line shows the actual temperatures, and the dotted lines show the 5% and 95% quantiles of the probabilistic forecast.*

of an ensemble forecast are equally likely and realistic, then the forecast can be used to create a probabilistic forecast relatively easily. A distribution can be fitted at each point in time, and dependencies between times can be analysed. If temperatures are assumed to be normally distributed, then fitting the distribution consists of calculating the mean and standard deviation at each lead time, and Pearson correlations can be used to measure dependence. If one is not content to assume normal distributions, then another parametric distribution, or possibly a non-parametric estimation method (such as kernel densities), can be used at each lead time. The measurement of dependencies becomes more difficult in this case, and there is no general framework under which all dependency structures can be parametrised. One simple and practical method is to assume that the dependencies can be captured well using Spearman correlations or Kendall's tau (Wang 1998).

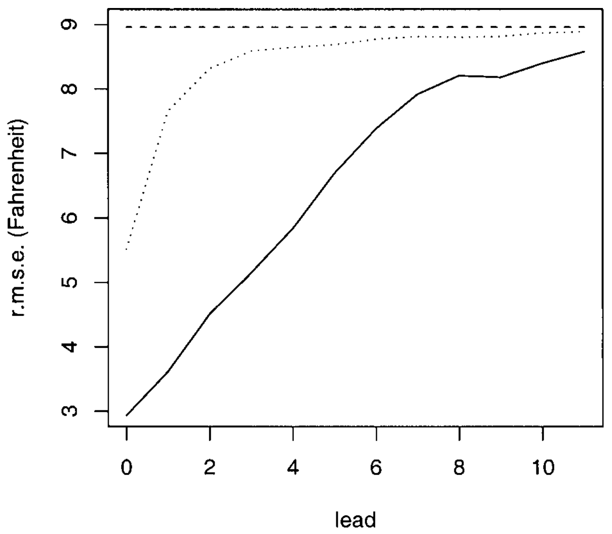
The assumption of different tracks in the ensemble being equally likely and realistic is not likely to be entirely correct, but errors can be partially accounted for by analysing past forecasts and using persistent biases in the distributions to adjust future forecasts (see Mylne et al. 2002, for an example of this approach). This will account for those errors in the ensemble that are not flow dependent. Flow-dependent errors will still remain, and can only be reduced by improving the underlying dynamical model.

The possible advantage of probabilistic forecasts based on ensembles rather than on single forecasts is that the variances and covariances of the forecast distribution can vary with the state of the atmosphere. Previous work, such as that by Denholm-Price & Mylne (2002), suggests that changes in the forecast variance that can be derived from ensembles may contain useful information beyond sampling variability. To the authors' knowledge, however, the accuracy of the predicted flow dependent *correlations* has never been tested, and so whether or not the predicted correlations (either in time or space) contain useful information is a matter of conjecture.

It would be possible to use the members of an ensemble directly to price weather derivatives, rather than by converting them to a probabilistic forecast. This is sometimes known as the 'end to end ensemble approach' (Palmer 2002). Each ensemble member could be used to calculate the outcome of the weather derivative, thus sampling the distribution of possible outcomes. There are a number of reasons, however, why the basic end to end method is not the most practical or accurate method for using the information in an ensemble forecast for weather derivative pricing (although it may be adequate for some other applications of ensemble forecasts). First, ensemble forecasts are generally much shorter than weather derivatives contracts: currently, the European Centre for Medium Range Weather Forecasting produces ensembles of length 10 days, and the National Centre for Environmental Prediction produces ensembles of length 16 days, while most weather derivative contracts are between one and five months in length. To overcome this, the ensemble could be used to estimate the distribution of outcomes for the first days of the contract (either 10 or 16 days) and historical data could be used to estimate the distribution of outcomes for the post-forecast period, with an assumption of independence between the two periods. But, as we shall see in Section 6.1, independence is a bad approximation in many cases, and as a result more complex methods must be used. Secondly, as mentioned above, it is generally necessary to correct the marginal distributions of temperature from ensemble forecasts. This can be done much more practically if a continuous distribution has been fitted to the ensemble. Thirdly, most ensemble forecasts have very small ensembles: at the time of writing the largest operational NWP ensemble is the ECMWF ensemble, with only 51 members. This limits the extent to which the probabilities of extreme events can be estimated. Fitting a continuous distribution allows the extrapolation of the forecast to extreme events. Such extrapolation may not be very accurate, but it is nevertheless essential for managers of weather derivative portfolios to have some estimate for the most extreme weather scenarios. The methods described in Section 7 below were developed in an attempt to overcome these three limitations of the direct application of ensemble forecasts, while still taking advantage of all the information contained in the forecast.

3.3. NWP versus stochastic model forecasts

Stochastic time-series models fitted to historical data from a given station can also be used to make forecasts: the predictability arises from the local autocorrelation



Q1

Figure 5. *RMS errors for single track temperature forecasts based on a dynamical model (solid line) and a local statistical model (dotted line). The dynamical model forecast shows dramatically smaller errors at all lead times.*

of temperature variability. In Caballero et al. (2002) (henceforth CJB) we showed that fractionally integrated auto-regressive moving average (ARFIMA) models (Granger & Joyeux 1980) capture the autocorrelation remarkably well. Such models are simple to implement and require trivial amounts of computer power, so they provide essentially cost-free forecasts. Thus, it is interesting to compare their forecast skill against that of full NWP models; if the comparison is favourable enough, it may be feasible to do without NWP altogether. (One could make the comparison quantitative through a cost/benefit analysis in which the benefit of not having to pay for NWP forecasts is balanced against the costs due to the lower skill of the stochastic model, but this is beyond the scope of the present study.) The use of stochastic model forecasts has been advocated by some authors (Diebold & Campbell 2001). Figure 5 shows the RMS error for NWP forecasts of Chicago temperatures, calculated using three months of actual forecasts. It also shows the RMS error for a prediction using an ARFIMA model for the same location, based on the assumption that the model is perfect (i.e. that the temperatures really were generated by this statistical model). We use a perfect model because it gives a reasonable estimate for the upper limit of predictability that could possibly be obtained using the stochastic model. We see that the stochastic model is significantly less accurate at all lead times. We conclude that there is little advantage is using the stochastic models for forecasting. Nonetheless, as we will see below, stochastic models can play an important role in derivative pricing even when NWP forecasts are used.

4. Pricing linear contracts on cumulative temperatures

The first kind of contract we consider will be a linear swap contract based on cumulative temperature. As we shall see, the use of forecasts to calculate the fair strike of such a contract is very simple. As we saw in Section 2, the fair strike is given by the expectation of the settlement index, $E(x)$. Applying the definition of cumulative temperatures (equation 3), and splitting the contract into historical, forecast and post-forecast periods (m, 12 and $M-m-12$ days long respectively), the fair strike on day m is given by:

$$E\left(\sum_{i=1}^{M} T_i\right) = \sum_{i=1}^{m-1} T_i^{hist} + \sum_{i=m}^{m+11} T_i^{cmf} + \sum_{i=m+12}^{M} T_i^{clim} \qquad (4)$$

where T_i^{hist} are historical temperatures, T_i^{cmf} are conditional mean temperature forecasts, and T_i^{clim} are climatological mean temperatures. Thus we see that the fair strike value can be estimated using only recent historical data, a conditional mean temperature forecast (most likely derived from an ensemble mean using regression) and climatological mean temperature values. Knowledge of moments of the distribution higher than the first is not needed in this simple case.

In order to go beyond calculating the fair strike value to calculating the distribution of outcomes of the linear swap on cumulative temperatures we have to determine the distribution of the cumulative temperature index. This in turn depends on the marginal distributions of temperatures on each day of the contract, and the dependencies between these distributions. Calculating the distribution of the index thus involves combining information from probabilistic forecasts with climatological models. The issues that need to be addressed when estimating this distribution are discussed in Section 6.1 and a method that can be used to make this combination in practice is described in Section 7.

5. Pricing linear contracts on degree days

We will now increase the complexity slightly by considering a linear swap based on heating degree days rather than cumulative temperatures. The fair strike on a particular day is still given by $E(x)$, the expectation of the settlement index x on that day. Denoting the daily values of the heating degree day index as y_i, we have:

$$E(x) = \sum_{i=1}^{M} E(y_i) \qquad (5)$$

If we assume that temperatures are normally distributed, then since y_i is a function of T_i, we can consider $E(y_i)$ to depend on the mean and standard deviation of T_i, which we denote by m_i and s_i. Given the definition of heating degree days (equation 1), this dependence can be evaluated rather easily. Letting f_i be the normal probability distribution function (PDF), with mean m_i and standard deviation s_i, and F_i its cumulative distribution function (CDF), we have:

$$
\begin{aligned}
E(y_i) &= \int_{-\infty}^{\infty} f_i(T)y(T)\,dT \\
&= \int_{-\infty}^{18} f_i(T)T\,dT \\
&= (18 - m_i)F_i(18) + s_i^2 f_i(18). \qquad (6)
\end{aligned}
$$

We then have:

$$
\begin{aligned}
E(x) = \sum_{i=1}^{m-1} \max\left(18 - T_i^{hist}, 0\right) + \sum_{i=m}^{m+11} \left(18 - m_i^{fc}\right)F_i(18) \\
+ s_i^{fc^2} f_i(18) + \sum_{i=m+12}^{M} \left(18 - m_i^{clim}\right)F_i(18) \\
+ s_i^{clim^2} f_i(18) \qquad (7)
\end{aligned}
$$

Thus we have written the fair swap strike in terms of historical data, T_i^{hist}, the means and variances of a probabilistic forecast, m_i^{fc} and $s_i^{fc^2}$ and the climatological temperature distribution, with mean and variance m_i^{clim} and $s_i^{clim^2}$.

Note that the correlations from the probabilistic forecast are not used. To the extent that the temperatures lie far below the baseline of 18°C used in the definition of heating degree days, the $F_i(18)$ will be close to 1 and the $f_i(18)$ will be close to zero. Equation 7 then reduces to equation 4. As with the linear index case, the distribution of outcomes of such contracts depends on the distribution of the index, estimation of which will be discussed in Sections 6.1 and 7.

6. Pricing non-linear contracts

We now consider the general situation in which the weather contract structure is non-linear. Of the contract types mentioned in the Introduction, this is relevant for capped swap contracts and option contracts. There are also many other non-linear contract structures, and indeed any shape is, in theory, possible.

For capped swap contracts, the first task is to estimate the fair strike, defined as the strike which gives an expected payoff of zero. In the special case in which the index distribution is close to normal and the caps are close to symmetrical with respect to the index distribution, the fair strike can be approximated as the expected value for the index, which can be calculated using the methods described for linear swaps using only conditional mean single track forecasts. When calculating the par value for the fair swap strike on standard indices, these approximations are often fairly good. However, during the contract period of pre-defined contracts it is highly unlikely that the caps will be symmetrical about the index distribution, since the caps are fixed in advance but the index distribution moves according to the latest weather data and forecasts. Calculating the fair swap strike is more difficult in this case and generally involves an iterative procedure based on an estimate of the whole distribution of the index x. Estimating the fair strike for the swap thus reduces to estimating this distribution.

For option contracts the fair price is given as the expectation of the distribution of payouts, $\mu_p = E(p(x))$. This too can only be evaluated using an estimate for the whole distribution of x.

We have seen that estimating the fair strike for swaps and the fair premium for options both reduce to estimating the distribution of the settlement index. How, then, can we estimate this distribution? For cumulative temperature indices based on normally distributed temperatures, the index will have a normal distribution because it is a sum of normal distributions. For heating degree day indices based on normally distributed temperatures the index distribution is not strictly a normal distribution because the temperature distributions are truncated by the definition of daily degree days (equation 1). However, in many cases, a normal distribution is a good approximation for a cumulative heating degree day index, perhaps because the contracts are often fairly long (30 to 150 days). For now we will thus assume that the index distribution can be represented by a normal distribution, although we only use this restriction to illustrate some important concepts: the methods we present for pricing can be used for non-normal distributions too, most conveniently by transforming data to a normal distribution at the beginning and transforming simulated results back to the original distribution at the end. Transformations that can be used for this purpose are discussed in Jewson & Caballero (2002).

Assuming a normal distribution for the settlement index, we can write the fair price for an option in terms of the mean μ and standard deviation σ of this distribution: $\mu_p = \mu_p(\mu, \sigma)$. Calculating the expected payoff thus reduces to estimating this mean μ and standard deviation σ. For a capped swap with index limits at L_1 and L_2, a tick of D, and a strike of S, Jewson (2003) shows that we can integrate the PDF of the index (which we are assuming is normal) against the definition of the payoff illustrated in Figure 1 to give:

$$
\begin{aligned}
\mu_p = D\bigg(\frac{1}{2}\delta L + \sigma^2(f(L_1) - f(L_2)) \\
+ F(L_1)(L_1 - \mu) + F(L_2)(\mu - L_2)\bigg) \qquad (8)
\end{aligned}
$$

where $\delta L = L_2 - L_1$.

For a call option with limit at L, strike at S and tick of D, with payoff illustrated in Figure 1 the analogous expression is:

$$\mu_p = D((L-S)+\sigma^2(f(S)-f(L))+F(L)(\mu-L) + F(S)(S-\mu)) \qquad (9)$$

This expression can be evaluated if we know the mean index μ and the standard deviation of the index σ. μ is given by the expressions we have already seen for the fair strike for a linear swap: equation 4 in the cumulative temperature case and equation 7 in the heating degree day case. The standard deviation of the index, σ, is, however, more difficult to estimate, and deriving an appropriate method for estimating the standard deviation of the index (and indeed, the whole index distribution) is the main purpose of Sections 6.1 and 7.

6.1. Autocorrelation and the index variance

We now discuss methods for estimating the standard deviation or variance and distribution of a weather index during the contract period. To put this in concrete terms, we might be estimating the distribution of possible values of the July mean temperature on 10 July, or estimating the probability of more than five more freezing days in December, on 5 December. We see from these examples that both forecasts and historical data will need to be used, as well as considerations of persistence of patterns of weather variability (i.e. autocorrelations).

First, we consider the standard deviation of a cumulative temperature index. The variance of a sum of random variables is the sum of the terms in the covariance matrix between the variables, and so:

$$\begin{aligned}\sigma^2 &= \sum_{i=1}^{M}\sum_{j=1}^{M} E(T_iT_j)\\ &= \sum_{i=m}^{m+11}\sum_{j=m}^{m+11} E(T_iT_j) + \sum_{i=m+12}^{M}\sum_{j=m+12}^{M} E(T_iT_j)\\ &\quad + 2\sum_{i=m}^{m+11}\sum_{j=M+12}^{M} E(T_iT_j) \qquad (10)\end{aligned}$$

The first of these terms represents covariances between temperatures during the forecast period. As we saw in Section 3, these are produced as part of a probabilistic forecast and can be calculated either from an ensemble forecast or from past forecast error statistics. The issue of how these two estimates are related is an interesting one which has not been addressed in the literature; doing so here would take us beyond the scope of this paper. For reasons of convenience, in our later examples we use past-forecast error statistics.

The second of these terms represents climatological temperature covariances, which can be estimated from historical data or from a daily temperature model fitted to historical data such as the ARFIMA model of CJB.

The third of these terms represents covariances between temperature during the forecast and the post-forecast period. This third term complicates matters significantly: if it were not for this term we could model the index variance as the sum of the index variance due to the forecast period and the index variance during the post-forecast period. However, the dependencies between these two periods, represented by the third term in equation 10, combined with the inherent positive autocorrelations of temperature, mean that this would always underestimate the total index variance.

An expression similar to equation 10, but slightly more complex, can be derived for heating degree day indices (not shown). The essence is the same: the total variance of the index depends not only on the variances of the index during the forecast and post-forecast periods, but also the covariances between these periods.

How important is this covariance term? If it is small, then perhaps it could be neglected without harm, and the modelling could be significantly simplified. We can estimate the size of the term by looking at the correlations between periods of temperature generated using a realistic statistical model. We fitted the CJB time series model to Chicago temperatures, and ran many simulations over periods of 90 days. We split the 90 days into a 12-day period to represent forecasts, and a post-forecast period with a length of between 1 day and 78 days. For each length of the post-forecast period we calculate the correlation between cumulative average temperature indices based on the simulated forecast period and the post-forecast period. The results are shown in Figure 6a. Not surprisingly, we see that the longer the post-forecast period, the lower the correlation between the two periods. We would expect most of this correlation to come from the long memory property of temperature variability that has been documented by CJB. To test this, and to test the likely values of this forecast/post-forecast correlation that could be seen at other stations, we vary the long memory parameter *d* while leaving the ARMA parameters with the values obtained from fitting to Chicago data. (The ARFIMA model use in CJB contains Q2
a parameter *d* that quantifies the rate of decay of the autocorrelations of temperature: higher values imply slower decay.) Figure 6 also shows the values for the correlation for various values of *d*. We see that most of the correlation does indeed depend on the presence of long memory, and that for stronger long memory (as is often seen in coastal regions of southern USA, for instance) the correlations between the two periods become large.

We now assess the size of this correlation effect in terms of the percentage underestimation of the total index variance that would result from assuming that these

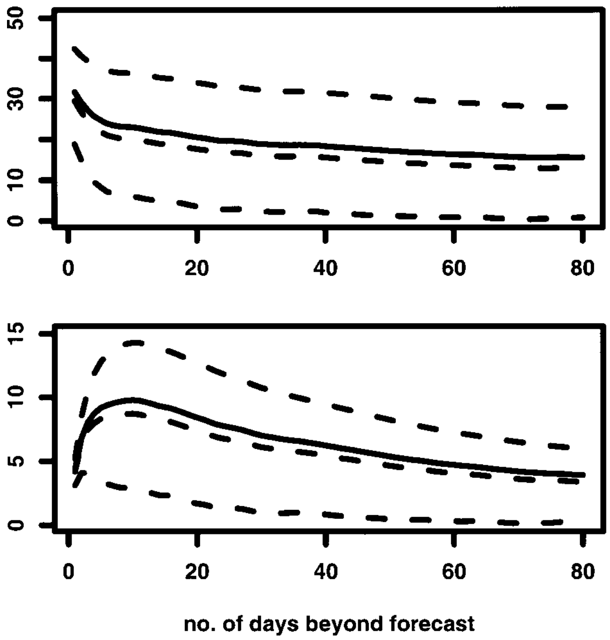


Figure 6. *The upper panel shows the correlation between cumulative temperature over a period of 11 days and a subsequent period of N days, where N is given on the horizontal axis. The correlations are based on simulations from the ARFIMA model of CJB fitted to Chicago temperatures (solid line) and on simulations based on the same model but with d readjusted to have the values 0.0, 0.1 and 0.2 (dashed lines, d increasing upward). The lower panel shows the error in the estimate of the standard deviation of cumulative temperature over a period of N + 11 days that is made if the first 11 days, and the subsequent N days, are assumed to be independent. The error is expressed as a percentage of the total actual standard of the whole period. Conventions as above.*

periods are independent. Figure 6b shows the results for the same range of values of d as used above. For each value of d we can see that the maximum error in the standard deviation would occur for a post-forecast period of length around 10 days, and has values of up to 20%. The maximum at 10 days can be understood from the expression for the total variance of the index. Letting σ_{fc}^2 be the variance for the forecast period and σ_{pfc}^2 be the variance for the post-forecast period, the error in the estimation of the total variance is given by:

$$\text{error} = \sigma^2 - \sigma_{fc}^2 - \sigma_{pfc}^2 = 2\rho\sigma_{fc}\sigma_{pfc} \qquad (11)$$

where ρ is the correlation between the forecast and post-forecast periods as shown in Figure 6a. We see that the error depends both on the correlation between the forecast and post-forecast period and also the size of the standard deviation of the post-forecast period. For post-forecast periods of much less than 10 days, the standard deviation of the post-forecast period is small, and so the error is small. For periods of much greater than 10 days, the correlation is small and so the error is small. For 10 days both the standard deviation and correlation have reasonable values, and the product of the two attains a maximum.

To summarise this section, we have shown that estimating the fair price of a non-linear weather contract involves estimating the distribution of the weather index on which the contract depends. Estimating the mean of this distribution is easy, but estimating the standard deviation is much harder since it depends on autocorrelations of temperature. The simplest model for estimating the standard deviation would be to assume independence between the forecast and post-forecast periods. However, we have shown that making such an assumption does not give an accurate estimate of the standard deviation of the index, and is particularly inaccurate when the contract extends around 10 days beyond the end of the forecast. We conclude that the assumption of independence between the forecast and post-forecast periods leads to significant inaccuracies and should be avoided if possible.

7. Methods for estimating the index variance

We have seen that a naive method for estimating the index variance based on assuming independence between the forecast and the post-forecast period is not accurate. We now present two methods which allow us to combine probabilistic forecasts and climatological models in such a way as to calculate accurate estimates of the distribution of the index without assuming independence between the forecast and post-forecast periods. Accurate estimates of the index distribution then allow us to calculate accurate prices for weather derivatives.

7.1. Pruning

The pruning method was described qualitatively by Jewson (2000) and Jewson et al. (2002a). It is based on the following steps:

- A daily time-series model for climatological temperatures is used to create a large ensemble of temperature tracks over the period of the contract.
- For the part of the time-series that lies during the forecast period, two probability densities are calculated for each track: one based on the probabilistic forecast distribution and the other based on the climatological distribution.
- The distribution of the index is estimated using contract indices based on each track, with weights for the track based on the ratio of the forecast density over the climatological density.

The mathematical rationale for this method (which was not given in the earlier articles) is the following: Let $p(\mathbf{T})$ be the payoff due to temperature track $\mathbf{T}$, $f(\mathbf{T})$ be the climatological probability of $\mathbf{T}$, and $g(\mathbf{T})$ be the forecast probability of $\mathbf{T}$. Then the climatological expected payoff μ_p^{clim} is given by:

$$\mu_p^{clim} = \int p(\mathbf{T}) f(\mathbf{T})\, d\mathbf{T} \tag{12}$$

where the integral is over all possible tracks for $\mathbf{T}$. The forecast expected payoff μ_p^{fc} is given by:

$$\mu_p^{fc} = \int p(\mathbf{T}) g(\mathbf{T})\, d\mathbf{T} \tag{13}$$

To evaluate μ_p^{clim} we choose a set of tracks that are equally spaced along the climatological CDF, $F(\mathbf{T})$. In other words, all the values of $dF(\mathbf{T}) = f(\mathbf{T})d\mathbf{T}$ are equal, so $dF(\mathbf{T}) = \frac{1}{N}$, where N is the number of tracks. The integral becomes:

$$\mu_p^{clim} = \int p(\mathbf{T})\, dF \tag{14}$$

$$\approx \frac{1}{N} \sum p(\mathbf{T}) \tag{15}$$

where the sum is over all the tracks in a discrete set of possible tracks.

If we evaluate μ_p^{fc} using the same set of tracks:

$$\mu_p^{fc} = \int p(\mathbf{T}) g(\mathbf{T})\, d\mathbf{T} \tag{16}$$

$$= \int p(\mathbf{T}) w(\mathbf{T})\, dF \tag{17}$$

$$\approx \frac{1}{N} \sum p(\mathbf{T}) w(\mathbf{T}) \tag{18}$$

where the weights $w(\mathbf{T})$ are given by:

$$w(\mathbf{T}) = \frac{g(\mathbf{T})}{f(\mathbf{T})} \tag{19}$$

In other words, we sum the payoffs for all possible tracks, but with weights. The weighting to be used is the forecast probability density of a certain track, divided by the climatological probability density of that track.

If we assume that both distributions are multivariate normal, then the densities can be evaluated using:

$$f(\mathbf{T}) = \frac{1}{(2\pi)^{\frac{d}{2}} \det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{T} - \mu)^t \Sigma^{-1} (\mathbf{T} - \mu) \right) \tag{20}$$

where Σ is the covariance matrix between temperatures on different days, and μ are the mean temperatures.

The advantages of this method for estimating the index distribution are that:

- it captures the correct correlations within the forecast period, within the climatological period, and between these two periods, and thus gives accurate estimates of the dispersion of the distribution;
- it allows easy calculation of both the par values and the new forecast-adjusted fair value; and
- it can be used to apply seasonal forecasts in exactly the same way as with weather forecasts.

7.2. Grafting

One disadvantage of pruning is that the simulations from the climatological model are not used very efficiently because some will have very low weights. For example, if the forecast is for warm with little probability of cold, then cold tracks will have low weights. An alternative method, which we call *grafting* involves sampling from the probabilistic forecast to create a large ensemble of tracks for the forecast period, and then extending these tracks using a climatological time-series model, with these resampled forecast tracks as initial conditions. It would also be possible to extend ensemble forecast tracks directly, but this would be less satisfactory since it would not sample extreme outcomes during the forecast period due to the small sizes of the ensembles available, and because bias correction of the forecast is more difficult to apply to the individual tracks than to the probabilistic forecast.

We illustrate the method by using the probabilistic forecast derived in Section 3 to price a 30-day option based on a non-linear temperature index. We calculate the fair value on the day the option starts.

First, we take the probabilistic forecast, and, assuming that the distribution of forecasts around the mean is given by the normal distribution, we generate 10,000 samples with the appropriate covariance structure. This can be done using standard numerical library routines. These 10,000 samples represent 10,000 possible outcomes for temperature during the forecast period. They contain the information from the forecast, but also extrapolate to extreme scenarios. If we have reason to believe that temperature at a particular location is not normally distributed, then we can apply transformations such as those described in Jewson & Caballero (2002) to convert to normal distributions.

Secondly, we use the ARFIMA model of CJB in forecast mode to extend these resampled forecasts out to the end of the contract. Because the temperature tracks for the first period are used as the initial conditions for the post-forecast period, covariances are modelled accurately between the two periods.

Thirdly, we convert each of the resulting 10,000 30-day simulated temperature tracks to index values, and determine the distribution for the settlement index. In the special case in which everything is assumed to be normally distributed and the index is a linear function of temperature, the distribution of the index can actually be derived without performing simulations, by summing values in the covariance matrix of daily temperatures

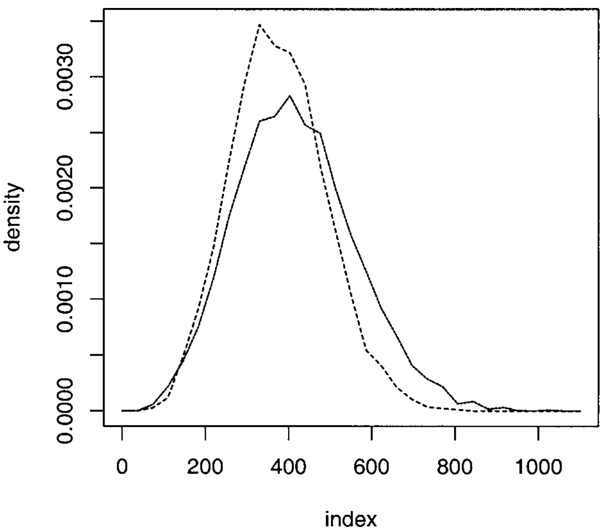


Figure 7. *Climatological distribution of a cumulative temperature index based on the ARFIMA model and 10,000 simulations (solid line). The dashed line shows the forecast distribution derived using the grafting method described in the text.*

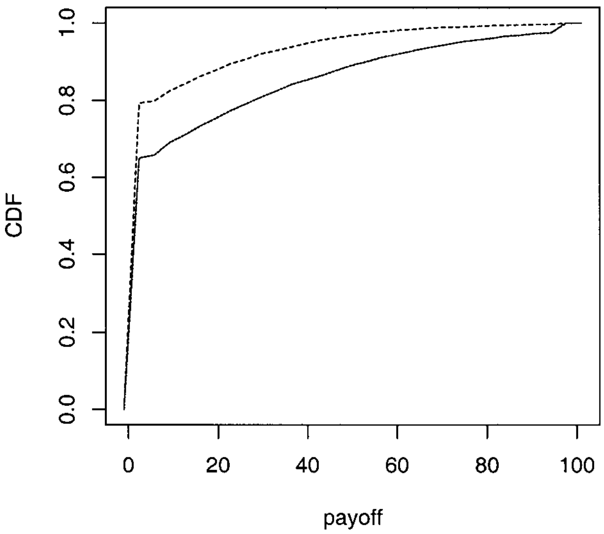


Figure 8. *The climatological and forecast CDFs for the payoff of an option, estimated using the climatological simulations from the ARFIMA model (solid line) and the grafting method described in the text (dashed).*

to get the index variance. In all other cases, where either temperatures are not assumed to be normally distributed, or the index definition is not a linear function of temperature, we have to use simulations. Since many contracts fall into this latter category, it is common to use simulations in all cases for simplicity and consistency.

Some results from applying this procedure are shown in Figure 7, which shows the index distribution derived only from climatological simulations, and from using the forecast shown in Figure 4 with the grafting method. The daily index is defined as $\max(T - 40, 0)$, and the total index is the sum of the daily indices. The daily index is strongly non-linear in the range of temperatures that occur during this period and so simulations must be used. We see that the addition of the forecast (which predicts colder than climatology on average over the forecast period) shifts the distribution of the settlement index to a lower and narrower range than predicted by climatology.

7.3. Payoff distributions

As we described in the introduction, sellers of weather derivatives need to estimate not only the fair value of each contract in their portfolio, but also the distribution of possible outcomes. It is particularly important to estimate the probabilities of extreme negative outcomes as well as possible. For both linear and non-linear weather contracts this involves estimating the distribution of the weather index. Pruning and grafting are both good methods for estimating this distribution since they are based on simulating temperatures directly and can be applied to any definition for the index and any contract structure.

The cumulative distributions for payoffs of an option based on the index described in the previous section are shown in Figure 8. The option has a strike of 450, a tick of 0.35 and a financial limit of 100. The effect of including the forecast can be seen to be to increase the probability of low payoffs of the option, as would be expected since the forecast is cold, and the option pays less for cold weather.

8. Summary

A number of studies have looked at quantitative applications of meteorological forecasts, such as Taylor & Buizza (2003), Smith et al.(2001) and Jewson & Ziehmann (2002). Taylor & Buizza (2003) and Smith et al. (2001) both describe 'end-to-end' application of ensemble forecasts, in which each ensemble member is converted directly into a variable of user interest. For Taylor & Buizza (2003) the variable of interest is electricity demand, while for Smith et al. (2001) it is bagel sales, electricity demand and wind power production.

In this paper we have discussed how weather forecasts can be used to estimate the fair prices and distributions of outcomes of weather derivatives contracts. This is important for contracts that have already started or that will start in the next 11 days. We have illustrated how both single and ensemble forecasts contain information that can be used in the pricing of different types of contract. We have argued that, in general, both single and ensemble forecasts should be converted into probabilistic forecasts to be most useful.

Our methods for using forecast information differ from the methods of Taylor & Buizza (2003) and Smith et al. (2001), in that we do not use the end to end approach. This is mainly because we are interested in the distributions of weather variables over periods longer than that covered by available forecasts. We are also interested in outcomes at more extreme levels than can be captured by the small ensembles currently used in ensemble forecasting systems. These two reasons lead us to take a number of approaches, including the pruning and grafting methods, that combine statistical climatological time-series models with forecast information.

The simplest case we consider is the estimation of the fair value of a linear (uncapped) swap contract based on cumulative average temperatures. In this case, only a single conditional mean weather forecast is needed. This forecast may be derived from an ensemble forecast, but probabilities of different events are not required.

The next most simple case is the estimation of the fair value of a linear contract based on degree days. Because of the non-linear definition of degree days with respect to temperature, we need a probabilistic forecast for this estimation. However, only the marginal distributions of the forecast are used, and not the dependency information.

The most general cases considered are a non-linear contract (a capped swap, or option) defined on either a cumulative temperature index or a degree day index, and the estimation of the distribution of outcomes of any contract, linear or non-linear. In these cases, probabilistic forecasts must be used, and both the marginal distributions and dependency information in the forecast are important. We have argued that ensemble forecasts cannot be used in a straightforward end-to-end approach for such valuation because, among other reasons, they are typically much shorter than the length of the weather contract. An end-to-end approach combined with independent historical data for the post-forecast part of the contract is also not accurate because it neglects important effects due to dependency between the weather in the forecast and post-forecast periods. We have evaluated the sizes of these effects in some simple cases, and shown that they cause a consistent underestimate of the variance of weather indices. We describe two methods that can be used to overcome this problem: pruning and grafting. They combine the best of climatological modelling of temperature (which captures the correct autocorrelations in time over the whole period of a weather contract) with the best of ensemble forecasts (which predict flow-dependent distributions over the first 12 days), and thus allow more accurate pricing of weather derivatives. We have not described the use of seasonal forecasts in detail, but note here that it should be possible to extend the use of the pruning methodology to such forecasts too.

Finally, we note that pruning and grafting are general methodologies that likely have applications beyond the pricing of weather derivatives. They give a solution to the general question of using the information from forecasts to predict the distribution of weather outcomes over a period longer than the period of the forecasts. Methods that assume independence between the forecast and post-forecast periods will always underestimate the dispersion of the distribution of weather events, because of the strong autocorrelations in weather variability. The greater the local autocorrelations in the weather, the greater this underestimation. Any application of forecasts that depends on correct estimation of the range of outcomes of weather over long periods should thus consider using these methodologies.

Acknowledgments

The authors would like to thank Anders Brix, Jeremy Penzer and Christine Ziehmann for useful discussion. Stephen Jewson is self-funded, while Rodrigo Caballero was supported by Danmarks Grundforskningsfond.

References

Caballero, R., Jewson, S. & Brix, A. (2002) Long memory in surface air temperature: detection, modelling and application to weather derivative valuation. *Climate Res.* **21**: 127–140.

Denholm-Price, J. & Mylne, K. (2002) Can an ensemble give anything more than gaussian probabilities? *Nonlinear Processes in Geophysics*. Submitted. **Q3, Q4**

Diebold, F. & Campbell, S. (2001) Weather forecasting for weather derivatives. University of Pennsylvania Institute for Economic Research, Tech. Rep. 8.

Freebairn, J. W. & Zillman, J. W. (2002) Economic benefits of meteorological services. *Meteorol. Appl.* **9**: 33–44.

Granger, C. W. J. & Joyeux, R. (1980) An introduction to long memory time series models and fractional differencing. *J. Time Ser. Anal.* **1**: 15–29.

Jewson, S. (2000) Use of GCM forecasts in financial-meteorological models. *Proceedings of the 25th Annual Climate Diagnostics and Prediction Workshop*, US Department of Commerce.

Jewson, S. (2003) Closed-form expressions for the pricing of weather derivatives: Part 1 – the expected payoff. *SSRN*. **Q5**

Jewson, S. & Caballero, R. (2002) Seasonality in the dynamics of surface air temperature and the pricing of weather derivatives. *Climate Res.* Submitted. **Q4**

Jewson, S. & Ziehmann, C. (2002) Weather swap pricing and the optimum size of medium range forecast ensembles. *Wea. and Forecasting*. **Q5**

Jewson, S., Brix, A. & Ziehmann, C. (2002a) Risk management. *Weather Risk Report*, Global Reinsurance Review, pp. 5–10.

Jewson, S., Brix, A. & Ziehmann, C. (2002b) Use of meteorological forecasts in weather derivative pricing. *Climate Risk and the Weather Market*, Risk Books, pp. 169–184. **Q6**

Leith, C. (1974) Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.* **102**: 409–418.

Mylne, K., Woolcock, C., Denholm-Price, J. & Darvell, R. (2002) Operational calibrated probability forecasts from the ECMWF ensemble prediction system: implementation and verification. *Preprints of the Symposium on Observations,* *Data Asimmilation and Probabilistic Prediction*, AMS, pp. 113–118.

Palmer, T. (2002) The economic value of ensemble forecasts as a tool for risk assessment: from days to decades. *Q. J. R. Meteorol. Soc.* **128**(581): 747–774.

Reiss, R. & Thomas, M. (1997) *Statistical Analysis of Extreme*
Q6 *Values*. Birkhauser.

Smith, L., Roulston, M. & von Hardenberg, J. (2001) End to end ensemble forecasting: towards evaluating the economic value of the ensemble prediction system. ECMWF, Tech. Rep.

Taylor, J. & Buizza, R. (2003) Using weather ensemble predictions in electricity demand forecasting. *Int. J. Forecasting* **19**: 57–70.

von Storch, H. & Zwiers, F. W. (1999) *Statistical Analysis in Climate Research*. Cambridge University Press.

Wang, S. (1998) Aggregation of correlated risk portfolios: models and algorithms. *Proc. Casualty Actuarial Soc.* **85**:
(163). Q7

Zeng, L. (2000) Weather derivatives and weather insurance: concept, application and analysis. *BAMS* **81**: 2075–2082.