View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

NCER
National Centre for Econometric Research

NCER Working Paper Series

Estimating the Payoffs of Temperature-based Weather Derivatives

A E Clements, A S Hurn, K A Liindsay

Working Paper #33

August 2008

Estimating the Payoffs of Temperature-based Weather Derivatives

A E Clements, A S Hurn
School of Economics and Finance, Queensland University of Technology

K A Lindsay
Department of Mathematics, University of Glasgow.

Abstract

Temperature-based weather derivatives are written on an index which is normally defined to be a nonlinear function of average daily temperatures. Recent empirical work has demonstrated the usefulness of simple time-series models of temperature for estimating the payoffs to these instruments. This paper argues that a more direct and parsimonious approach is to model the time-series behaviour of the index itself, provided a sufficiently rich supply of historical data is available. A data set comprising average daily temperature spanning over a hundred years for four Australian cities is assembled. The data is then used to compare the actual payoffs of temperature-based European call options with the expected payoffs computed from historical temperature records and two time-series approaches. It is concluded that expected payoffs computed directly from historical records perform poorly by comparison with the expected payoffs generated by means of competing time-series models. It is also found that modeling the relevant temperature index directly is superior to modeling average daily temperatures.

Keywords
Temperature, Weather Derivatives, Cooling Degree Days, Time-series Models.

JEL Classification Numbers
C14, C52.

Corresponding author
Adam Clements
School of Economics and Finance
Queensland University of Technology
Brisbane, 4001
Qld, Australia

email `a.clements@qut.edu.au`

1 Introduction

There has been growing interest in weather derivatives that permit the financial risk associated with climatic conditions such as temperature or rainfall to be managed. Similar to the situation in financial markets where a derivative security takes its value from an underlying financial asset or index, a weather derivative takes its value from an underlying measure of weather, such as temperature, rainfall or snowfall over a particular period of time. The first weather derivative was transacted in the US in 1996 and the size of the market is now in excess of US$ 8 billion.[1] Because temperature and precipitation intrinsically cannot be traded, there is no arbitrage-free pricing framework available to price these weather derivatives. Consequently this paper is primarily concerned with the development of accurate estimates of the expected payoffs of weather derivatives which is the crux of any pricing strategy.

Despite the existence of precipitation-based derivatives, the vast majority of all weather derivatives are based on a temperature indices, such as heating degree days and cooling degree days.[2] Temperature derivatives are currently written on temperature indices collected from several US and European cities as well as two Japanese cities. Major participants in this market include utilities and insurance companies along with other firms with costs or revenues that are dependent upon temperature. For example, an electricity supplier normally provides its customers with electricity at a fixed price irrespective of the wholesale price. On the other hand the wholesale price of electricity can fluctuate wildly with extreme temperatures, and so temperature-based derivatives can provide a hedging tool for fluctuations in wholesale electricity prices. Consequently the focus of this paper will be exclusively on temperature-based derivatives.

Various methods for estimating expected payoffs have been suggested. The most straightforward of these computes the expected value of payoffs from historical records (Zeng, 2000; Platen and West, 2003). In isolated cases and for very simple models, closed form solutions for the expected payoff can be developed (Benth and Šaltytė-Benth, 2005). A more general method is to fit a model to the time-series of average temperature so as to capture seasonal variations in both temperature and its volatility (Platen and West, 2003; Campbell and Diebold, 2004). The model is then used to simulate temperature outcomes over the period of the contract in order to construct the distribution of the temperature-based index on which the derivative is written. Note that widely-available meteorological forecasts are not suitable for this purpose because these forecasts are made over relatively short horizons, such as 7 days, whereas temperature derivatives are often traded well before[3] contracts generate any payoffs (Wilks, 1995; Jewson and Caballero, 2003; Campbell and Diebold, 2004).

This paper argues that a more direct and parsimonious way of approximating the distribution of the temperature-based index on which derivatives are written is to build a time-series model of the index itself, rather than a model of underlying average daily temperature. It is further

[1]The first recorded activity was an over-the-counter heating degree day swap option between Entergy-Koch and Enron for the winter of 1997 in Milwaukee, Wisconsin (Tindall 2006).

[2]Garmen *et al.*, 2000 posit that 98-99% of all weather derivatives currently traded are based on temperature.

[3]For example, participants in temperature-based weather derivative may enter into a contract many months before the arrival of the summer on which the payoff of the contract is to be determined.

shown that this approach gives rise to a simple yet effective model with empirical performance which is superior to methods based on historical payoffs and models of average temperature. Of course, this approach requires an rich supply of temperature data on which to build the time-series of the index.

For the empirical work in this paper a data set comprising average daily temperatures for over a century for four Australian cities, namely, Brisbane, Melbourne, Perth and Sydney was collected. These locations were chosen primarily because they are the four major cities of Australia, and also because accurate temperature records of long-duration are available at single weather stations, an important institutional requirement for writing temperature-based derivatives. This is a quality data set which represents a substantial improvement on what appears to be the current standard used in the literature. The potential downside of using Australian temperature data is that Australia currently has no organised market for temperature derivatives such as that organised by the Chicago Mercantile Exchange (CME) or the London International Financial Futures and Options Exchange (Liffe).[4] Consequently, no actually observed derivative prices can be used in this analysis. Nevertheless, the methodology developed here is generally applicable and could be used to estimate the payoffs to temperature derivatives in any market.

The rest of the paper is structured as follows. Section 2 outlines the concept of the 'tick value' of a temperature-based derivative and the importance of expected payoff in its pricing. Section 3 describes the data used in this investigation. Section 4 presents a seasonal GARCH model of average daily temperature as developed by Campbell and Diebold (2004) and Section 5 builds the time-series model of the actual index on which the derivatives are written. Alternative approaches to modeling the distribution of the payoffs from temperature-based call options, in the form of expected profit, are evaluated in Section 6. Section 7 is a brief conclusion.

2 Tick Values of Temperature Options

The most commonly referenced weather indices on which temperature derivatives are written are heating degree days (HDDs) and cooling degree days (CDDs). Let $T^{\max}$ and $T^{\min}$ be respectively the maximum and minimum temperatures in degrees Celsius measured on a particular day at a specific weather station. The HDD and CDD indices at that station on that day are defined respectively by

$$\begin{aligned} \text{HDD} &= \max\big(0,\, 18 - T\big)\,, \\ \text{CDD} &= \max\big(0,\, T - 18\big)\,, \end{aligned} \tag{1}$$

where T is the arithmetic mean of the maximum and minimum temperatures achieved on that day, namely

$$T = \frac{T^{\max} + T^{\min}}{2}\,. \tag{2}$$

The choice of threshold, in this instance 18°C, is set by market convention and is the standard used in the US. In the southern (northern) hemisphere the HDD (CDD) season would be from

[4]Trading of weather derivatives on the CME began in September 1999 and by 2006 approximately 55% of all weather derivative trading was transacted on the CME. By contrast, in 2004 Liffe started trading weather derivatives in July 2001 but suspended trading in these instruments in 2004 due to a lack of turnover (Tindall 2006).

May to September, while the CDD (HDD) season would be from November to March.

Temperature-based call options are based on *cumulative* heating or cooling degree days constructed by summing daily HDD/CDD indices over a period of N days to get

$$\begin{aligned} \mathrm{H}_N &= \sum_{k=1}^{N} \max\big(0,\, 18 - T_k\big)\,, \\ \mathrm{C}_N &= \sum_{k=1}^{N} \max\big(0,\, T_k - 18\big)\,, \end{aligned} \tag{3}$$

where T_k is the mean temperature, defined as in equation (2), on the kth day of the life of the option. Without loss of generality, the analysis of this paper will be limited to considering European call options written on CDDs. The choice of European option is not limiting in the sense that many more complex derivative strategies are in fact combinations of simple European options. The choice of CDDs is more pragmatic, driven by the fact that CDDs are uniformly important to all the major Australian cities in the data set.

Let D be the strike price of a temperature based option defined as a particular value of the relevant cumulative index. The buyer of a vanilla European call option pays an up-front premium and receives a payout if the value of the relevant index exceeds the strike price, D, at the maturity of the option. The tick value of an CDD call option with strike price D and duration N days is therefore

$$\mathcal{T}_N = \max\big(\mathrm{C}_N - D, 0\big)\,. \tag{4}$$

The actual monetary payoff from the contract is the product of the tick value and the tick size, defined as the cash value of a tick.

Traditionally, the valuation of options under schemes such as that of Black and Scholes (1973) discounts the expected payoff at the riskless force of interest. This choice of discount rate is based on a zero-arbitrage argument involving the formation of a portfolio consisting of a riskless combination of an option and the underlying asset. However, in context of a temperature-based weather derivative, the underlying indices are not tradable, and therefore these derivatives cannot be priced by means of a zero-arbitrage argument.

The most common practical approach used to price temperature-based derivatives is the *actuarial* valuation method, discussed, for example, in Zeng (2000) and Platen and West (2003). Broadly speaking this approach prices the derivative at the mean expected payout plus a premium for overhead expense. The simplest way of implementing this pricing scheme is to review historical records of C_N over the period of a contract in previous years and use these values to calculate the hypothetical payout of the contract had it been in place. The actuarially fair price for the derivative would then be the mean historical payoff.

Recent empirical work by Campbell and Diebold (2004) fits a time-series model to daily temperature to capture seasonal climatic patterns. The resulting model is then used to simulate the probability density function, $f(x)$, and associated cumulative distribution function, $F(x)$, of the relevant cumulative index over the period of the contract. In the case of a call option for N days with strike price D, for example, the expected tick value of the payoff is

$$\mathrm{E}\,[\,\mathcal{T}_N\,] = \int_D^\infty (x - D)\, f(x)\, dx = \int_D^\infty \big(1 - F(x)\big)\, dx\,.$$

This approach will be the subject matter of Section 4. An alternative approach proposed in this paper will be to model the CDD index itself, rather than model the underlying temperature from which the CDDs are subsequently calculated. This direct approach which is likely to be less prone to error than the dealing with temperature is outlined in Section 5.

3 Data

The data set comprises daily maximum and minimum temperature records in degrees Celsius for Brisbane, Melbourne, Perth and Sydney.[5] Following standard practice in pricing weather derivatives (Zeng, 2000; Platen and West, 2003; and Campbell and Diebold, 2004), the analysis is conducted on the time series of average daily temperatures computed as the arithmetic mean of the daily maximum and minimum values, as in equation (2). For all the data sets, instances of single missing values were treated by averaging adjacent records. In a few rare cases where several days were missing, the long term average for those days was inserted. Finally, following Campbell and Diebold (2004), all occurrences of the 29 February were removed.

Brisbane, Melbourne, Perth and Sydney were chosen primarily because they are the four major cities of Australia, and also because accurate temperature records of over 100 years are available for these cities at comparable weather stations. The construction of the temperature record for each city is now discussed in more detail.

Brisbane The temperature record contains 44043 observations starting on the 1/1/1887 and ending on 31/8/2007. The time series is constructed from data collected from three weather stations: Brisbane Regional Office (Station Number 40214) 1/1/1887 - 31/3/1986; Brisbane Airport (Station Number 40223) 1/4/1986 - 14/2/2000); and again from Brisbane Airport (Station Number 40842) 15/2/2000 - 31/8/2007.

Melbourne The temperature record contains 55358 observations starting on 1/1/1856 and ending on 31/8/2007. The time series is a continuous set of observations made at the Melbourne Regional Office (Station Number 86071) weather station. The location of the office changed in the early 1980s although the name of station did not.

Perth The temperature record contains 40393 observations starting on 1/1/1897 and ending on 31/8/2007. The time series is constructed from data collected at two weather stations: Perth Regional Office (Station Number 9034) 1/1/1897 - 2/6/1944; and Perth Airport (Station Number 9021) 3/6/1944 - 31/8/2007.

Sydney The temperature record contains 54263 observations starting on 1/1/1859 and ending on 31/8/2007. The time series is a continuous set of observations made at the Sydney Observatory Hill (Station Number 66062) weather station.

[5] All the raw data were supplied by Climate Information Services, National Climate Centre, Australian Bureau of Meteorology.

Summary statistics for the average daily temperatures are reported in Table 1. Brisbane is the hottest city on average and also records the lowest variability in average daily temperature. Melbourne is the coldest on average and has a relatively high variability in average daily temperature. Perth has the most variable daily temperatures. There are significant differences in all the cities between the sample means of temperature pre- and post-1950. This suggests that a time trend will be an important component of a model of average daily temperatures.[6] Interestingly, any trend in daily temperatures seems to be driven by the increasing minimum value of daily temperatures rather than by an increasing maximum value.

Summary Statistics							
	Dates	N	Mean	Med.	S. Dev.	Max.	Min
Brisbane	1887 - 2006	43800	20.52	20.85	4.05	34.65	8.30
Brisbane	1887 - 1949	22995	20.39	20.70	4.11	34.65	8.30
Brisbane	1950 - 2006	20805	20.67	21.00	3.97	34.15	8.45
Melbourne	1856 - 2006	55115	14.95	14.40	4.74	34.55	2.25
Melbourne	1856 - 1949	34310	14.64	14.15	4.72	34.20	2.25
Melbourne	1950 - 2006	20805	15.46	14.90	4.72	34.55	3.80
Perth	1897 - 2006	40150	18.07	17.25	4.94	36.95	6.25
Perth	1897 - 1949	19345	17.92	17.20	4.72	36.95	6.25
Perth	1950 - 2006	20805	18.21	17.25	5.15	36.80	6.25
Sydney	1859 - 2006	54020	17.66	17.80	4.28	33.75	6.40
Sydney	1859 - 1949	33215	17.34	17.50	4.32	33.70	6.40
Sydney	1950 - 2006	20805	18.18	18.25	4.15	33.75	7.70

Table 1: Mean, median, standard deviation, maximum and minimum of average daily temperature in four Australian cities. Note that the sample is curtailed to end on 31 December 2006 to ensure that summary statistics are computed over complete years.

Figures 1 and 2 show the long-term expected values and standard deviations of daily temperatures for each day of the year. Figure 1 shows that all the cities have similar seasonal fluctuation and that the estimates of the long-term expected values of temperature on each day in every city is converging. By contrast, Figure 2 demonstrates more variability in the seasonal pattern of the volatility of temperatures across the cities. It is also noticeable that, despite the length of the temperature records, the estimates of daily volatility appear not to have converged to the same extent as the estimates of the mean temperature.

[6]Given the location of the actual weather stations from which the time-series data are assembled, it is conjectured that this time trend is probably due to urbanisation rather than a manifestation of global warming.

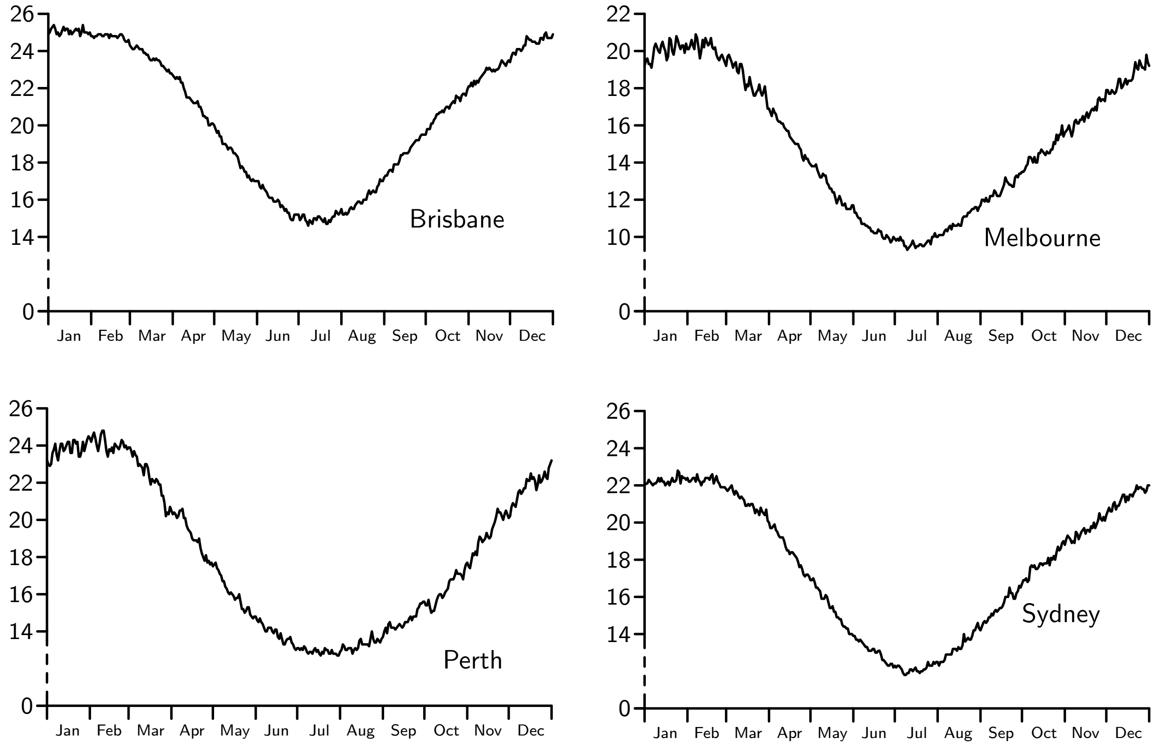

Figure 1: Plots of the expected value of the average daily temperatures for the four Australian cities.

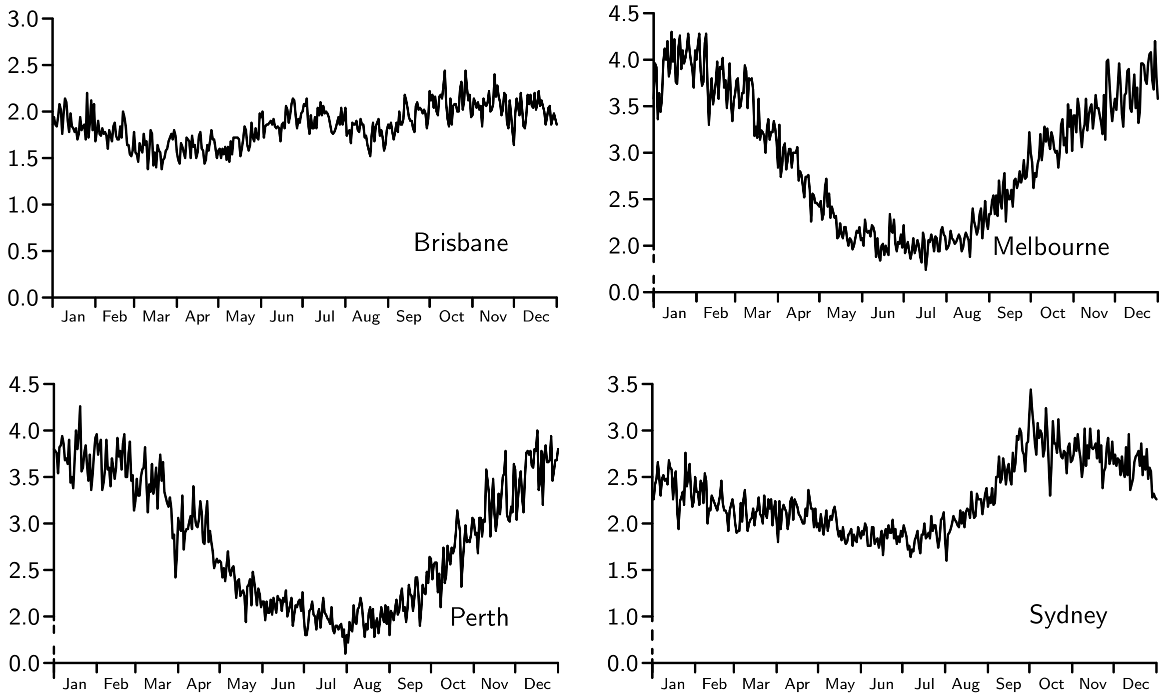

Figure 2: Plots of the expected value of the volatility of average daily temperatures for the four Australian cities.

4 Modeling Temperature

The model of this section will follow in spirit the analysis of Campbell and Diebold (2004), namely, to see if simple time-series models, similar in structure for each city, can provide an adequate model of temperature. If so, then repeated simulation of the model will allow accurate pricing of temperature-based weather derivatives. The details of the implementation differ from those in Campbell and Diebold (2004) and are now described.

For all cities, the temperature, T_t, is the average daily temperature defined in equation (2). Following the general convention (Davis, 2001, Alaton *et al.*, 2002, Benth and Šaltytė-Benth, 2005), the deviations of temperature from its long-term average $\theta_t = T_t - \bar{T}_t$ are modeled as a low-order autoregressive (AR) process[7]

$$T_t = \bar{T}_t + \sum_{j=1}^{m} \alpha_j \, \theta_{t-j} + \sigma_t \, \varepsilon_t \, , \tag{5}$$

where m is the order of the AR process, ε_t is an $iid(0\,,1)$ process and $\bar{T}_t$ is modelled as the sum of a trend and a periodic component by the expression

$$\bar{T}_t = \gamma_0 + \gamma_1 \, \mathrm{Trend}_t + \sum_{j=1}^{k} \phi_j \cos\left(\frac{2\pi jt}{365}\right) + \sum_{j=1}^{k} \varphi_j \sin\left(\frac{2\pi jt}{365}\right) . \tag{6}$$

In order to capture both the observed seasonal pattern (see Figure 2) of the volatility of temperature and any persistence in volatility, a conventional GARCH(1,1) model (Bollerslev, 1986) is augmented by adding a constant seasonal component as a forcing variable in the conditional variance equation, as in Campbell and Diebold (2004). The SGARCH(1,1) model[8] for conditional variance is then given by

$$\sigma_t^2 = \beta_0 + \beta_1 \, \sigma_{t-1}^2 + \beta_2 \, \varepsilon_{t-1}^2 + \beta_3 \, S_t \tag{7}$$

where

$$S_t = \xi_0 + \sum_{j=1}^{m} \xi_j \cos\left(\frac{2\pi jt}{365}\right) + \sum_{j=1}^{m} \delta_j \sin\left(\frac{2\pi jt}{365}\right) . \tag{8}$$

Table 2 reports the estimation results for the SGARCH(1,1) model of temperature for the entire sample period. To reduce the dimension of the optimisation problem, the parameters of equation (6) and equation (8) are pre-computed by ordinary least squares using temperature and squared deviations of temperature from $\bar{T}_t$ as the dependent variables respectively.

[7] Alternatively, a fractionally integrated process for deviations could be used (see, for example, Caballero and Jewson, 2002), but this modeling avenue is not pursued here.

[8] Experimentation with other models of conditional variance, for example, where volatility is given by a persistent and transitory component in the spirit of the component-GARCH, model of Engle and Lee (1993), suggested that the SGARCH(1,1) model was a satisfactory way to model the seasonal level in conditional variance.

	Cities			
	Brisbane	Melbourne	Perth	Sydney
α_0	−0.0066	−0.0835	−0.0561	0.0175
	(0.0063)	(0.0091)	(0.0097)	(0.0074)
α_1	0.6857	0.6487	0.7715	0.6264
	(0.0053)	(0.0047)	(0.0054)	(0.0049)
α_2	−0.0695	−0.1332	−0.1646	−0.0753
	(0.0064)	(0.0058)	(0.0069)	(0.0055)
α_3	0.0180	−0.0045	−0.0164	0.0263
	(0.0060)	(0.0055)	(0.0065)	(0.0050)
α_4	−0.0022	0.0050	−0.0035	−0.0178
	(0.0057)	(0.0052)	(0.0063)	(0.0047)
α_5	0.0184	0.0113	0.0158	0.0094
	(0.0057)	(0.0051)	(0.0062)	(0.0046)
α_6	0.0187	0.0049	0.0028	0.0148
	(0.0055)	(0.0050)	(0.0061)	(0.0046)
α_7	0.0000	−0.0062	0.0136	0.0097
	(0.0045)	(0.0041)	(0.0049)	(0.0039)
β_0	0.3544	0.7329	1.4262	0.7722
	(0.0201)	(0.0452)	(0.0979)	(0.0259)
β_1	0.1545	0.1126	0.0980	0.1842
	(0.0057)	(0.0051)	(0.0063)	(0.0056)
β_2	0.4209	0.5160	0.2290	0.2137
	(0.0186)	(0.0242)	(0.0469)	(0.0167)
β_3	0.0419	0.0159	0.0191	0.0497
	(0.0021)	(0.0010)	(0.0014)	(0.0015)

Table 2: Parameter estimates for AR(7) - SGARCH(1,1) models for average daily temperature data.

The results conform largely with expectations although there are a few features worthy of comment. The AR(1) parameter is particularly strong while the coefficients of the other lagged deviations of temperature are smaller, but mostly statistically significant. Experimentation with a longer lag structure, as used by Campbell and Diebold (2004), did not significantly alter the main thrust of the results.[9] The coefficients of the variance equation are all significant including, β_3, the coefficient on the exogenous seasonal pattern in the conditional variance equation. Interestingly enough, the inclusion of this term seems to dampen the estimate of persistence in volatility by comparison with the kinds of estimates usually obtained in GARCH models of financial asset returns, where the sum $\beta_1 + \beta_2$ is typically very close to 1.

The emphasis in this paper is on the use of a generic model which takes a common structure across all cities. Despite this simple modeling strategy, plots of the standardised residuals from the AR(7)-SGARCH(1,1) model, illustrated in Figure 3, suggest that these residuals are approximately standard normal. In Brisbane and Sydney the residuals appear slightly more peaked than the standard normal and Melbourne exhibits a marginal skew. On the whole, however, it

[9] While this procedure certainly kills any autocorrelation in the residuals for the lag lengths used, the real problem in modeling temperature is that limited structure remains out to very long lag lengths.

may be argued that the model does an adequate job in capturing the main characteristics of the dynamics of average temperatures in the major cities.

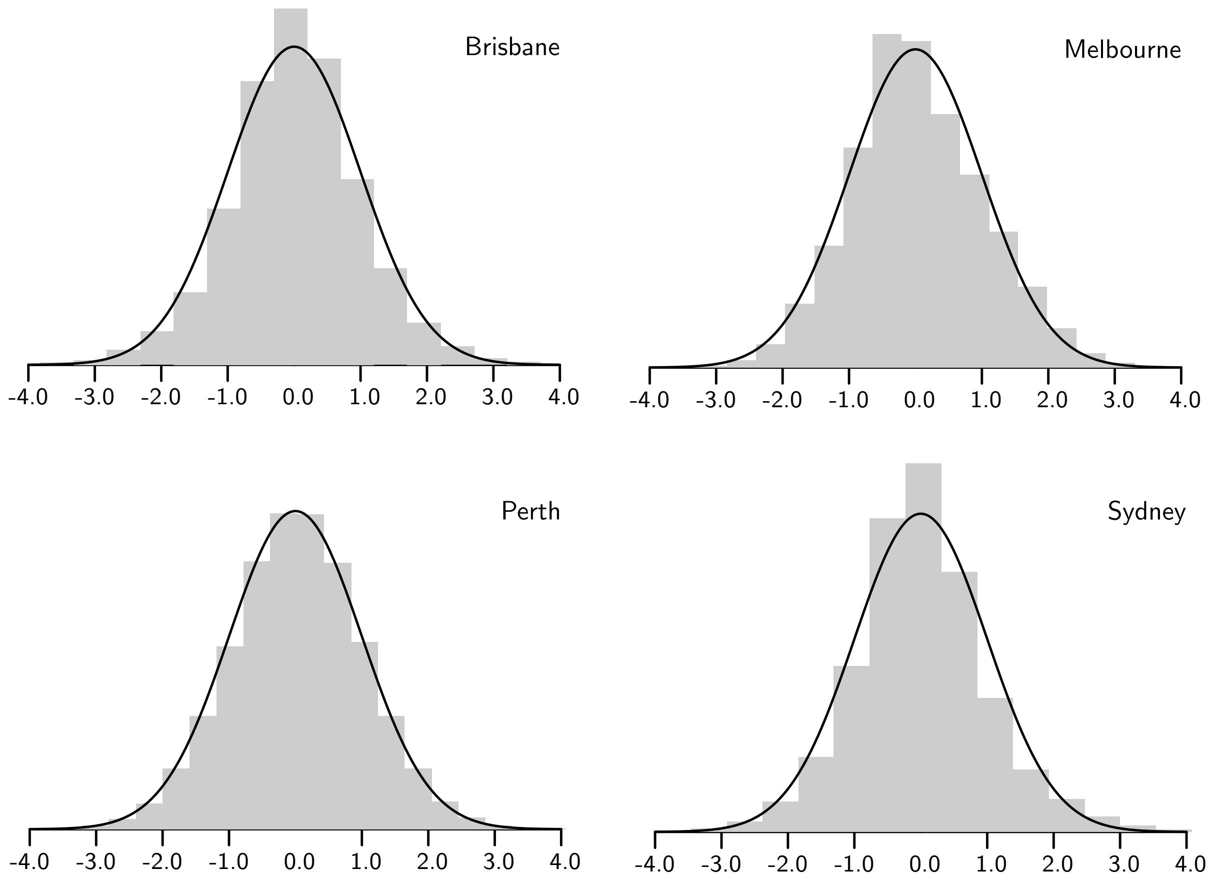

Figure 3: Histograms of standardised residuals from the $SGARCH(1,1)$ models with a standard normal curve superimposed.

Given the parameter estimates for the AR(7)-SGARCH(1,1) model, equations (5) and (7) may be used to simulate realisations of average daily temperatures. From a series of k simulations, realisations $\widehat{C}_1, \widehat{C}_2, \cdots, \widehat{C}_k$ of cumulative CDDs for the appropriate period may be obtained. These are to be regarded as k independent drawings from the distribution of cumulative CDDs for the period under consideration. Given a strike price D, the j-th tick value for the realised cumulative CDD $\widehat{C}_j$ is

$$\mathcal{T}_N = \left[\begin{array}{cc} 0 & \widehat{C}_j \le D\,, \\ \widehat{C}_j - D & \widehat{C}_j > D \end{array} \right.$$

The expected tick value is then given by the Monte Carlo estimate

$$\mathrm{E}\,[\,\mathcal{T}_N\,] = \int_D^\infty (x-D) f(x)\,dx \approx \frac{1}{k}\sum_{j=1}^{k} (\widehat{C}_j - D) H(\widehat{C}_j - D)\,, \tag{9}$$

where $H(\cdot)$ is the Heaviside function defined by

$$H(x) = \left[\begin{array}{cc} 1 & x > 0\,, \\ 0 & x \le 0\,. \end{array} \right.$$

5 Modeling Cumulative Degree Days

An alternative and more straightforward approach to evaluating the expected tick value of a temperature derivative contract is to model cumulative CDDs directly, on the assumption that a sufficiently long temperature record exists to make this feasible. Let $C_1, C_2, \cdots, C_n$ be time series of n historical observations of cumulative CDDs. Intuitively, in the absence of a trend in the temperature data from which the cumulative CDDs are derived, these observations will be independently and identically distributed realisations from the distribution of cumulative CDDs. A simple quadratic trend model is proposed for cumulative CDDs. Although the quadratic term is not expected to be significant, it is included to account for the possibility of piecewise trends in cumulative CDDs due to the effect of urbanisation late in the sample period. Accordingly, cumulative CDDs are described by the general model

$$C_t = \eta_0 + \eta_1 \text{ Trend}_t + \eta_2 \text{ Trend}_t^2 + \epsilon_t$$

where ϵ_t is now distributed $iid(0, \sigma_\epsilon^2)$.

Estimation of the parameters of this model for each city yields

$$\begin{aligned}
\mathrm{E}[C_t]_{\text{Brisbane}} &= \underset{(14.8071)}{564.0290} + \underset{(0.5603)}{0.2617} \text{ Trend}_{\mathrm{t}} + \underset{(0.0044)}{0.0008} \text{ Trend}_t^2 \\
\mathrm{E}[C_t]_{\text{Melbourne}} &= \underset{(14.1135)}{192.4003} - \underset{(0.4259)}{0.5819} \text{ Trend}_{\mathrm{t}} + \underset{(0.0027)}{0.0077} \text{ Trend}_t^2 \\
\mathrm{E}[C_t]_{\text{Perth}} &= \underset{(19.7863)}{410.2985} + \underset{(0.8155)}{1.2290} \text{ Trend}_{\mathrm{t}} + \underset{(0.0071)}{0.0025} \text{ Trend}_t^2 \\
\mathrm{E}[C_t]_{\text{Sydney}} &= \underset{(11.6784)}{311.1655} - \underset{(0.3595)}{0.1276} \text{ Trend}_{\mathrm{t}} + \underset{(0.0023)}{0.0065} \text{ Trend}_t^2
\end{aligned}$$

where the figures in parentheses are standard errors. The quadratic terms are significant in Melbourne and Sydney, while Perth contains a linear trend. The effect is less marked in Brisbane, although there does seem to be a small linear trend effect. These results are reinforced by the time-series plots of cumulative CDDs in Figure 4.

The actual cumulative CDD at time $t = n + 1$ is

$$\mathrm{E}[C_{n+1}] = \widehat{\eta}_0 + \widehat{\eta}_1 \text{ Trend}_{n+1} + \widehat{\eta}_2 \text{ Trend}_{n+1}^2$$

and the actual cumulative CDD at time $t = n + 1$ may be regarded as $\mathrm{E}[C_{n+1}] + \widehat{\epsilon}$ where $\widehat{\epsilon}$ is a draw from the distribution of ϵ. On the assumption that ϵ_t is iid, a nonparametric kernel may be used to estimate the probability density function of the disturbances from

$$\widehat{f}(\epsilon) = \frac{1}{nh} \sum_{k=1}^{n} K\Big(\frac{\epsilon - \widehat{\epsilon}_k}{h}\Big)$$

based on the observed regression residuals, $\widehat{\epsilon}_1, \widehat{\epsilon}_2, \cdots, \widehat{\epsilon}_n$, given that $K(\cdot)$ is the Gaussian kernel function and h is the kernel bandwidth.

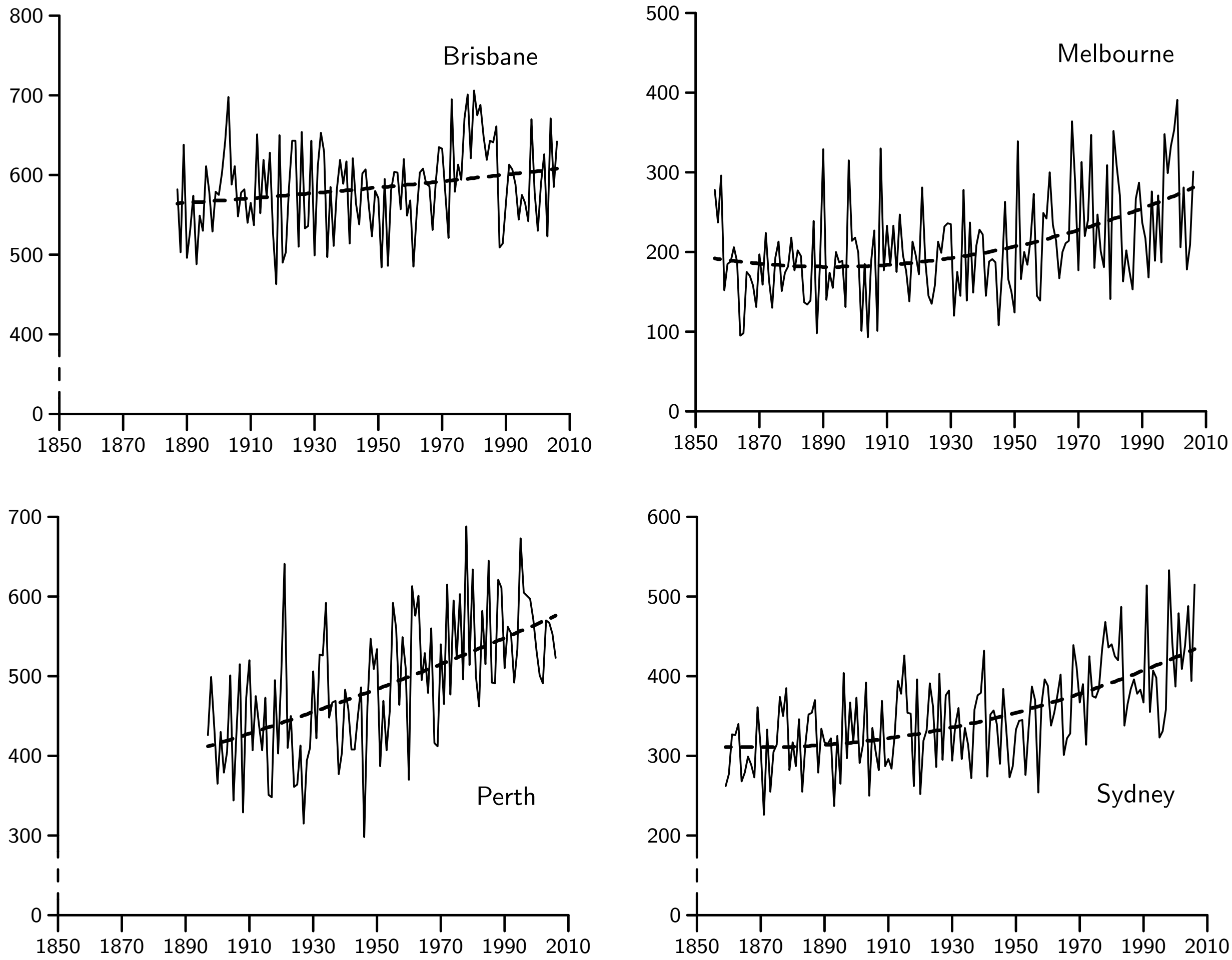

Figure 4: Time series of cumulative CDDs for each city with estimated trend component superimposed (dashed line).

Although the time series model of cumulative CDDs is simpler than the AR(7)-SGARCH(1,1) model of daily temperature, the construction of the expected tick value of the generic call option is slightly more difficult in this instance. At time $t = n+1$, the tick value of the call option, with strike price D, is

$$\mathcal{T}_N = \left[\begin{array}{cc} 0 & \mathrm{E}[C_{n+1}] + \epsilon \le D \\ \mathrm{E}[C_{n+1}] + \epsilon - D & \mathrm{E}[C_{n+1}] + \epsilon > D \end{array} \right.$$

The expected tick value is then given by

$$\begin{aligned} \int_D^\infty (x-D) f(x)\, dx &= \int_{D-\mathrm{E}[C_{n+1}]}^\infty (\mathrm{E}[C_{n+1}] + \epsilon - D) \widehat{f}(\epsilon) d\epsilon \\ &= \frac{1}{nh} \sum_{k=1}^n \int_{D-\mathrm{E}[C_{n+1}]}^\infty \big(\mathrm{E}[C_{n+1}] + \epsilon - D\big) K\Big(\frac{\epsilon - \hat{\epsilon}_k}{h}\Big) d\epsilon \end{aligned}$$

where $x = \mathrm{E}[C_{n+1}] + \epsilon$ and the density function of x is identical to that of ϵ. The change of variable from ϵ to $y = (\epsilon - \widehat{\epsilon}_k)/h$ yields

$$\mathrm{E}[\mathcal{T}_N] = \frac{1}{n} \sum_{k=1}^n \int_{\frac{\mathrm{E}[C_{n+1}] + \widehat{\epsilon}_k - D}{h}}^\infty (\mathrm{E}[C_{n+1}] + \widehat{\epsilon}_k + hy - D) K(y)\, dy\,.$$

It is convenient to write $z_k = (D - \mathrm{E}[C_{n+1}] - \widehat{\epsilon}_k)/h$ so that

$$\mathrm{E}[\mathcal{T}_N] = \frac{h}{n} \sum_{k=1}^{n} \int_{z_k}^{\infty} (y - z_k) K(y)\, dy$$

When $K(\cdot)$ is taken to be the Gaussian kernel, then

$$\begin{aligned} \int_{z_k}^{\infty} y K(y)\, dy &= \phi(z_k) \\ \int_{z_k}^{\infty} z_k K(y)\, dy &= z_k(1 - \Phi(z_k)) \end{aligned}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal PDF and CDF respectively. It follows immediately that the expected tick value takes the simple form

$$\mathrm{E}[\mathcal{T}_N] = \frac{h}{n} \sum_{k=1}^{n} \left[\phi(z_k) - z_k(1 - \Phi(z_k))\right] . \tag{10}$$

6 Computing Expected Payoffs

Before presenting the comparison of expected payoffs associated with the methods suggested in Sections 4 and 5 it is instructive to look at the distributions of cumulative CDDs. Descriptive statistics for cumulative CDDs are reported in Table 3 and their distributions, in terms of histograms, are plotted in Figure 5.

	Summary Statistics					
	N	Mean	Med.	S. Dev.	Max.	Min
Brisbane	121	584.2	584.6	54.49	463.3	705.9
Melbourne	152	207.9	195.6	64.09	93.5	391.4
Perth	111	489.6	492.2	83.30	298.3	688.3
Sydney	149	350.0	350.2	60.07	225.5	533.3

Table 3: Mean, median, standard deviation, maximum and minimum cumulative CDDs in four Australian cities.

The descriptive statistics for cumulative CDDs are very much as expected given the geographical location of the cities. There are, however, two observations of note arising out of Table 3. It is apparent that the distribution of cumulative CDDs for Melbourne is skewed to the right as evidenced by a mean which is significantly larger than the median. This is to be expected given both the instances of extreme heat in Melbourne and the strength of the trend in the Melbourne CDD data identified in Section 5. Perth, on the other hand, is notable for the diffuse nature of the distribution of cumulative CDDs, recording a standard deviation significantly larger than those of all of the other cities. These features of the distributions are also apparent from the histograms in Figure 5. The histograms for Brisbane and Perth appear to be symmetrical differing only in the range of CDDs. The histogram for Sydney, however, suggests that there may be a slight skew in favour of higher CDDs, but not as pronounced as that of Melbourne. At first sight, therefore, these distributions look well behaved and could be taken as reasonable evidence in favour of using historical records to price temperature-based derivatives. As will become apparent, however, these marginal distributions mask the fact that CDDs are strongly correlated over time.

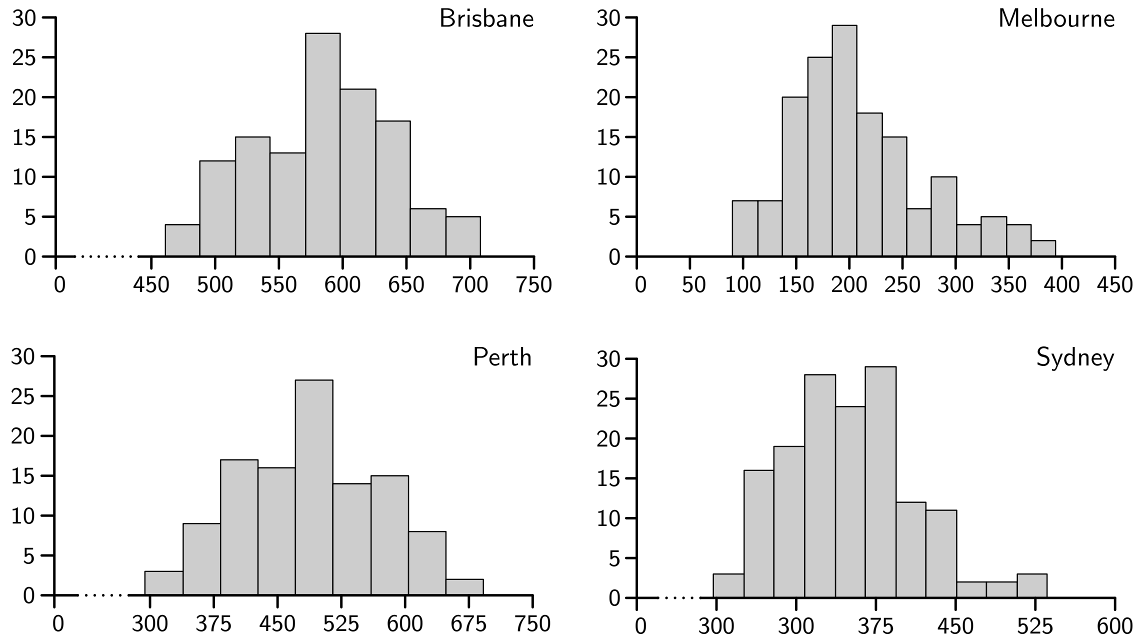

Figure 5: Histograms of cumulative CDDs for each city.

The task is now to provide a means of comparing each method of establishing the expected tick value of a temperature call option, namely, historical temperature records, the AR(7)-SGARCH model and the simple time-series model of CDDs. In this paper, the metric for comparison is taken to be the mean 'profit' of the call option contracts over a period of years, where profit is defined to be the difference between the actual tick value of the contract and the expected tick value or 'price' of the option. Of course, this is not meant to represent a true price for the option as this notional pricing strategy takes no account of discounting or overhead expenses. But of course any pricing scheme will stand or fall by its ability to estimate the expected tick value accurately.

The profits of two separate call option contracts written on the period 1 January to 31 March are reported in Tables 4 and 5 respectively. The experiments begin by pricing these options for the year 1950 using data up to and including 1949. The actual payoff for 1950 is recorded, the profit or loss stored and the data set is updated to include all the temperature records for the next year. These steps are repeated up to and including 2007 giving a total of 58 separate profits for each option. The call options used in the experiment have respective strike prices $D = \mu + 0.5\sigma$ and $D = \mu + 0.75\sigma$ where μ is the unconditional mean and σ is the unconditional standard deviation of CDDs up to the current year under consideration. The means and standard deviations of the profits are regarded as measures of the performance of each method for determining expected tick values.

	Brisbane	Melbourne	Perth	Sydney
Historical				
Mean Payoff	−12.9194	−9.2412	−18.5485	15.9323
SDev Payoff	30.1388	49.2567	48.2584	51.2786
SGARCH(1,1)				
Mean Payoff	1.8261	17.5827	8.3870	32.6240
SDev Payoff	30.1484	48.4349	47.0758	51.5130
CumCDD				
Mean Payoff	2.0639	0.2028	−5.6663	0.4997
SDev Payoff	30.2010	47.5819	48.3131	50.2641

Table 4: Means and standard deviations of payoffs to temperature 90-day call option defined on CDDs with strike price $D = \mu + 0.5\sigma$, where μ and σ are the unconditional mean and standard deviation of available historical CDDs. The option is priced for each year from 1950 to 2007 inclusive.

	Brisbane	Melbourne	Perth	Sydney
Historical				
Mean Profit	−17.1256	−20.7096	−26.0408	7.0359
SDev Profit	24.9354	42.9552	40.0616	47.9302
SGARCH(1,1)				
Mean Profit	−1.9947	11.7679	0.0267	23.7435
SDev Profit	24.6708	42.7484	39.9373	48.1691
CumCDD				
Mean Profit	1.6206	1.4344	−5.5060	0.6781
SDev Profit	24.8074	41.8817	41.1706	47.1369

Table 5: Means and standard deviations of profits to 90-day call options defined on CDDs with strike price $D = \mu + 0.75\sigma$, where μ and σ are the unconditional mean and standard deviation of available historical CDDs. The option is priced for each year from 1950 to 2007 inclusive.

A number of conclusions emerge from the results reported in Tables 4 and 5. The most important conclusion is that historical pricing is the least robust of all the procedures for estimating payoffs, particularly for call options with higher strike prices. The historical method seems to underprice significantly in 3 of the 4 cities. It is conjectured that this failure is due to ignoring the trend in cumulative CDDs in combination with the fact that the resolution of the empirical distribution will be particularly grainy when the data are relatively sparse. The SGARCH(1,1) model performs with credit relative to the historical method, but nevertheless there appears to be a general tendency to overprice the European call option particularly for Melbourne and Sydney. Furthermore, the estimated payoffs based on the SGARCH(1,1) model appears not to be degraded when call options with higher strike prices are considered. It is clear, however, that the method of estimating payoffs based on modeling the time-series of cumulative CDDs directly offers superior performance relative to the other two methods. The mean of estimated payoffs from this approach is estimated relatively accurately, with the standard deviation of the distribution being data dependent and therefore almost identical for each method.

7 Conclusion

This paper has compared three methods for estimating the tick value of European call options written on cooling degree days, a temperature-based index constructed as a nonlinear function of average daily temperature. Although the cooling degree day index is the focus of this research, the methods used are equally applicable to derivatives based on heating degree days. The conclusions reached in this investigation may be succinctly summarised as follows. Historical methods for estimating the tick value of these options appear to be unreliable and are to be treated with some scepticism. If too long a run of data is used to estimate payoffs, any trends in the index are likely to be under-emphasised. On the other hand, looking at only recent data is likely to cause problems in terms of the poor resolution of the distribution of the relevant index. Consequently, model-based estimation of tick values is to be preferred. Moreover, attempts to model the relevant index directly are likely to be simpler, more parsimonious and ultimately more successful at estimating payoffs than more complex models that are built on average daily temperatures.

References

Alaton, P., Djehiche, B. and Stillberger, D. (2002). On Modelling and Pricing Weather Derivatives, *Applied Mathematical Finance*, **9**, 1-20.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, **31**, 307-327.

Benth, F.E. and Šaltytė-Benth, J. (2005). Stochastic Modelling of Temperature Variations with a View Toward Weather Derivatives, *Applied Mathematical Finance*, **12**, 53-85.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**, 637-659.

Caballero, R., Jewson S. and Brix, A. (2002). Lon Memory in Surface Air Temperature Detection, Modelin and Application to Weather Derivative Valuation, *Climate Reasearch*, **21**, 127-140.

Campbell and Diebold. (2004). Weather Forecasting for Weather Derivatives. *Journal of the American Statistical Society*, **100**, 6-16.

Davis M.H.A. (2001). Pricing Weather Derivatives by Marginal Value, *Quantitative Finance*, **1**, 1-4.

Engle, R.F. and Lee, G.G.J. (1993). A Permanent and Transitory Component Model of Stock Return Volatility, University of California, San Diego, Discussion Paper 92-44R, November 1993.

Garman, M., Blanco, C. and Erickson, R. (2000). Weather Derivatives: Instruments and Pricing Issues. *Environmental Finance*, March.

Jewson S. and Caballero, R. (2003). The Use of Weather Forecasts in the Pricing of Weather Dervivatives, *Meterological Applications*, **10**, 377-389.

Platen, E. and West, J. (2003). Fair Pricing of Weather Derivatives, Quantitative Finance Research Centre, University of Technology, Research Paper Series, 106.

Tindall, J. (2006). Weather Derivatives: Pricing and Risk Management Applications. *Unpublished manuscript*, Institute of Actuaries of Australia.

Wilks, D.S. (1995). Statistical Methods in the Atmospheric Sciences. New York: Academic Press.

Zeng, L. (2000). Pricing Weather Derivatives, *Journal of Risk Finance*, **Spring**, 72-78.