Presented in the 11th International Conference on Engineering Applications of Neural Networks (EANN). London, UK (2009).

Modeling and Forecasting CAT and HDD Indices for Weather Derivative Pricing

A. Zapranis[1], A. Alexandridis[1*]

[1]Department of Accounting and Finance, University of Macedonia of Economics and Social Studies, 156 Egnatia St., P.O. 54006, Thessaloniki, Greece.
zapranis@uom.gr, alex@uom.gr

Abstract. In this paper, we use wavelet neural networks in order to model a mean-reverting Ornstein-Uhlenbeck temperature process, with seasonality in the level and volatility. We forecast up to two months ahead out of sample daily temperatures and we simulate the corresponding Cumulative Average Temperature and Heating Degree Day indices. The proposed model is validated in 8 European and 5 USA cities all traded in Chicago Mercantile Exchange. Our results suggest that the proposed method outperforms alternative pricing methods proposed in prior studies in most cases. Our findings suggest that wavelet networks can model the temperature process very well and consequently they constitute a very accurate and efficient tool for weather derivatives pricing. Finally, we provide the pricing equations for temperature futures on Heating Degree Day index.

Keywords: Weather Derivatives, Pricing, Forecasting, Wavelet Networks

1 Introduction

Recently a new class of financial instruments -weather derivatives- has been introduced. The purpose of weather derivatives is to allow business to insure themselves against fluctuations in the weather. According to [1, 2] nearly $1 trillion of the US economy is directly exposed to weather risk. Just as traditional contingent claims, whose payoffs depend upon the price of some fundamental, a weather derivative has its underlying measure such as: rainfall, temperature, humidity or snowfall. Weather derivatives are used to hedge volume risk, rather than price risk,

The Chicago Mercantile Exchange (CME) reports that the estimated value of its weather products reached $22 billion through September 2005, with more than 600,000 contracts traded. This represents sharp rise comparing with 2004 in which notional value was $2.2 billion [3]. Moreover, it is anticipated that the weather market will continue to develop, broadening its scope in terms of geography, client base and

*Corresponding author: e-mail: aalex@uom.gr

inter-relationship with other financial and insurance markets. In order to fully exploit all the advantages that this market offers, adequate pricing approach is required.

Weather risk is unique in that it is highly localized, and despite great advances in meteorological science, still cannot be predicted precisely and consistently. Weather derivatives are also different than other financial derivatives in that the underlying weather index, like Heating Degree Days (HDD), Cooling Degree Days (CDD), Cumulative Average Temperature (CAT), etc. cannot be traded. Mathematical explanations of these indices are given in section 4. Furthermore, the corresponding market is relatively illiquid. Consequently, since weather derivatives cannot be cost-efficiently replicated with other weather derivatives, arbitrage pricing cannot directly apply to them. The weather derivatives market is a classic incomplete market, because the underlying weather variables are not tradable

The first and simplest method that has been used in weather derivative pricing is burn analysis. Burn analysis is just a simple calculation of how a weather derivative would perform in the past years. By taking the average of these values an estimate of the price of the derivative is obtained. Burn analysis is very easy in calculation since there is no need to fit the distribution of the temperature or to solve any stochastic differential equation. Moreover, burn analysis is based in very few assumptions. First, we have to assume that the temperature time-series is stationary. Next, we have to assume that the data for different years are independent and identically distributed. For a detailed explanation of Burn Analysis we refer to [4].

A closer inspection of a temperature time series shows that none of these assumptions are correct. It is clear that the temperature time-series is not stationary since it contains seasonalities, jumps and trends [5, 6]. Also the independence of the years is under question. [4] show that these assumptions can be used if the data can be cleaned and detrended. Although their results show that the pricing still remains inaccurate. Other methods as index and daily modeling are more accurate but still burn analysis is usually a good first approximation of the derivative's price.

In contrast to the previous methods, a dynamic model can be used which directly simulates the future behavior of temperature. Using models for daily temperatures can, in principle, lead to more accurate pricing than modeling temperature indices. On the other hand, deriving an accurate model for the daily temperature is not a straightforward process. Observed temperatures show seasonality in all of the mean, variance, distribution and autocorrelations and long memory in the autocorrelations. The risk with daily modeling is that small misspecifications in the models can lead to large mispricing in the contracts.

The continuous processes used for modeling daily temperatures usually take a mean-reverting form, which has to be descretized in order to estimate its various parameters. Previous works suggest that the parameter of the speed of mean reversion, α, is constant. The work of [5] indicates exactly the opposite. In addition, α is important for the correct and accurate pricing of temperature derivatives, [7]. In [5] the parameter a was modeled by a Neural Network (NN). More precisely wavelet analysis was used in order to identify the trend and the seasonal part of the temperature signal and then a NN was used in the detrended and deseasonalized series. However WA is limited to applications of small input dimension since its constructing wavelet basis of a large input dimension is computationally expensive, [8]. In NN framework the initial values of the weights are randomly chosen which usually leads to large training times and to convergence to a local minimum of a specified loss function. Finally, the use of sigmoid NN does not provide any information about the network construction. In this study we expand the work of [5] by combining these two steps. To overcome these problems we use networks with wavelets as activation functions, namely Wavelet Networks, (WN). More precisely we use truncated Fourier series to remove the seasonality and the seasonal volatility of the temperature in various locations as in [9]. Next a wavelet network is constructed in order to fit the daily average temperature in 13 cities and to forecast daily temperature up to two months hoping that the waveform of the activation functions of the feedforward network will fit much better the temperature process than the classical sigmoid functions. For a concise treatment of wavelet analysis we refer to [10-12] while for wavelet networks we refer to the works of [8, 13, 14]. The forecasting accuracy from the proposed methodology is validated in a two month ahead out of sample window. More precisely the proposed methodology is compared against historical burn analysis and the Benth & Benth's model in forecasting CAT and HDD indices. Finally, we extend the work of [5] by presenting the pricing equations for future HDD contracts when the speed of mean reversion is not constant.

The rest of the paper is organized as follows. In section 2, we describe the process used to model the average daily temperature. In section 3 a brief introduction to wavelet networks is given. In section 4 we describe our data and apply our model to real data. In section 5 we discuss CAT and HDD derivatives pricing and finally, in section 6 we conclude

2 Modeling Temperature Process

Many different models have been proposed in order to describe the dynamics of a temperature process. Early models were using AR(1) processes or continuous equivalents [7, 15, 16]. Others like [17] and [18] have suggested versions of a more general ARMA(p,q) model. [19] have shown, however, that all these models fail to capture the slow time decay of the autocorrelations of temperature and hence lead to significant underpricing of weather options. Thus more complex models were proposed, like an Ornstein-Uhlenbeck process [20]. Also in the noise part of the process, the Brownian noise was at first replaced by a fractional Brownian noise and then by a Levy process [21]. A temperature Ornstein-Uhlenbeck process is:

$$dT(t) = dS(t) - \kappa(T(t) - S(t))dt + \sigma(t)dB(t) \tag{1}$$

where, $T(t)$ is the daily average temperature, $B(t)$ is a standard Brownian motion, $S(t)$ is a deterministic function modelling the trend and seasonality of the average temperature, while $\sigma(t)$ is the daily volatility of temperature variations. In [9] both $S(t)$ and $\sigma^2(t)$ were modeled as truncated Fourier series:

$$S(t) = a + bt + a_0 + \sum_{i=1}^{I_1} a_i \sin(2i\pi(t - f_i)/365) + \sum_{j=1}^{J_1} b_j \cos(2j\pi(t - g_j)/365) \quad (2)$$

$$\sigma^2(t) = c + \sum_{i=1}^{I_2} c_i \sin(2i\pi t/365) + \sum_{j=1}^{J_2} d_j \cos(2j\pi t/365) \quad (3)$$

From the Ito formula an explicit solution for (1) can be derived:

$$T(t) = s(t) + (T(t-1) - s(t-1))e^{-\kappa t} + \int_{t-1}^{t} \sigma(u) e^{-\kappa(t-u)} dB(u) \quad (4)$$

According to this representation *T(t)* is normally distributed at *t* and it is reverting to a mean defined by *S(t)*. A discrete approximation to the Ito formula, (4), which is the solution to the mean reverting Ornstein-Uhlenbeck process (1), is:

$$T(t+1) - T(t) = S(t+1) - S(t) - (1 - e^{-k})\{T(t) - S(t)\} + \sigma(t)\{B(t+1) - B(t)\} \quad (5)$$

which can be written as:

$$\tilde{T}(t+1) = a\tilde{T}(t) + \tilde{\sigma}(t)\varepsilon(t) \quad (6)$$

where

$$\tilde{T}(t) = T(t) - S(t) \quad (7)$$

$$a = e^{-k} \quad (8)$$

In order to estimate model (6) we need first to remove the trend and seasonality components from the average temperature series. The trend and the seasonality of daily average temperatures is modeled and removed as in [9]. Next a wavelet neural network is used to model and forecast daily detrended and deseasonalized temperatures. Hence, equation (6) reduces to:

$$T(t) = \varphi\big(T(t-1)\big) + e_t \quad (9)$$

where $\varphi(\bullet)$ is estimated non-parametrically by a wavelet network. Hence the parameter α is not constant. Once we have the estimator of the underlying function φ, then we can compute the daily values of α as follows:

$$a = d\tilde{T}(t+1)/d\tilde{T}(t) = d\varphi / d\tilde{T} \quad (10)$$

The analytic expression for the wavelet network derivative $d\varphi / d\tilde{T}$ can be found in [14]. Due to space limitations we will refer to the works of [5, 9] for the estimation of parameters in equations (2), (3), (6) and (8).

3 Wavelet Neural Networks for Multivariate Process Modeling.

Here the emphasis is in presenting the theory and mathematics of wavelet neural networks. So far in literature various structures of a WN have been proposed [8, 13, 22]. In this study we use a multidimensional wavelet neural network with a linear connection of the wavelons to the output. Moreover in order for the model to perform well in linear cases we use direct connections from the input layer to the output layer. A network with zero hidden units (HU) is the linear model.

The network output is given by the following expression:

$$\hat{y}(\mathbf{x}) = w_{\lambda+1}^{[2]} + \sum_{j=1}^{\lambda} w_j^{[2]} \cdot \Psi_j(\mathbf{x}) + \sum_{i=1}^{m} w_i^{[0]} \cdot x_i \tag{11}$$

In that expression, $\Psi_j(\mathbf{x})$ is a multidimensional wavelet which is constructed by the product of m scalar wavelets, $\mathbf{x}$ is the input vector, m is the number of network inputs, λ is the number of hidden units and w stands for a network weight. Following [23] we use as a mother wavelet the Mexican Hat function. The multidimensional wavelets are computed as follows:

$$\Psi_j(\mathbf{x}) = \prod_{i=1}^{m} \psi(z_{ij}) \tag{12}$$

where ψ is the mother wavelet and

$$z_{ij} \frac{x_i - w_{(\xi)ij}^{[1]}}{w_{(\zeta)ij}^{[1]}} \tag{13}$$

In the above expression, $i = 1, \ldots, m$, $j = 1, \ldots, \lambda+1$ and the weights w correspond to the translation ($w_{(\xi)ij}^{[1]}$) and the dilation ($w_{(\zeta)ij}^{[1]}$) factors. The complete vector of the network parameters comprises:

$$w = \left(w_i^{[0]}, w_j^{[2]}, w_{\lambda+1}^{[2]}, w_{(\xi)ij}^{[1]}, w_{(\zeta)ij}^{[1]} \right) \tag{14}$$

There are several approaches to train a WN. In our implementation we have used ordinary back-propagation which is less fast but also less prone to sensitivity to initial conditions than higher order alternatives. The weights $w_i^{[0]}$, $w_j^{[2]}$ and parameters $w_{(\xi)ij}^{[1]}$ and $w_{(\zeta)ij}^{[1]}$ are trained for approximating the target function.

In WN, in contrast to NN that use sigmoid functions, selecting initial values of the dilation and translation parameters randomly may not be suitable, [22]. A wavelet is a waveform of effectively limited duration that has an average value of zero and localized properties hence a random initialization may lead to wavelons with a value of zero. Also random initialization affects the speed of training and may lead to a local minimum of the loss function, [24]. In literature more complex initialization methods have been proposed, [13, 23, 25]. All methods can be summed in the following three steps.

1. Construct a library W of wavelets
2. Remove the wavelets that their support does not contain any sample points of the training data.
3. Rank the remaining wavelets and select the best regressors.

The wavelet library can be constructed either by an orthogonal wavelet or a wavelet frame. However orthogonal wavelets cannot be expressed in closed form. It is shown that a family of compactly supported non-orthogonal wavelets is more appropriate for function approximation, [26]. The wavelet library may contain a large number of wavelets. In practice it is impossible to count infinite frame or basis terms. However arbitrary truncations may lead to large errors, [27].

In [23] three alternative methods were proposed in order to reduce and rank the wavelet in the wavelet library namely the Residual Based Selection (RBS) a Stepwise Selection by Orthogonalization (SSO) and a Backward Elimination (BE) algorithm. In this study we use the BE initialization method that proved in previous studies to outperform the other two methods, [14, 23].

All the above methods are used just for the initialization of the dilation and translation parameters. Then the network is further trained in order to obtain the vector of the parameters $w = w_0$ which minimizes the cost function.

4 Modeling and Forecasting CAT and HDD indices

In this section real weather data will be used in order to validate our model. The data set consists of 4015 values, corresponding to the average daily temperatures of 11 years (1995-2005) in Paris, Stockholm, Rome, Madrid, Barcelona, Amsterdam, London and Oslo in Europe and New York, Atlanta, Chicago, Portland and Philadelphia in USA. Derivatives on the above cities are traded in CME. In order for each year to have equal observations the 29th of February was removed from the data. Finally, the model was validated in data consisting of 2 months, January – February, of daily average temperatures (2005-2006) corresponding to 59 values. Note that meteorological forecasts over 10 days are not considered accurate.

The list of traded contracts in weather derivatives market is extensive and constantly evolving. However over 90% of the contracts are written on temperature CAT and HDD indices. In Europe, CME weather contracts for the summer months are based on an index of CAT. The CAT index is the sum of the daily average temperatures over the contract period. The average temperature is measured as the simple average of the minimum and maximum temperature over one day. The value of a CAT index for the time interval $[\tau_1, \tau_2]$ is given by the following expression:

$$\int_{\tau_1}^{\tau_2} T(s)ds \tag{15}$$

where the temperature is measured in degrees of Celsius. In USA, CME weather derivatives are based on HDD or CDD index. A HDD is the number of degrees by which daily temperature is below a base temperature, while a CDD is the number of degrees by which the daily temperature is above the base temperature,

i.e., Daily HDD = max (0, base temperature – daily average temperature),

Daily CDD = max (0, daily average temperature – base temperature).

The base temperature is usually 65 degrees Fahrenheit in the US and 18 degrees Celsius in Europe. HDDs and CDDs are usually accumulated over a month or over a season. At the end of 2008, at CME were traded weather derivatives for 24 US cities[2], 10 European cities[3], 2 Japanese cities[4] and 6 Canadian cities[5].

Table 1 shows the descriptive statistics of the temperature in each city for the past 11 years. The mean and standard deviation HDD represent the mean and the standard deviation of the HDD index for the past 11 years for a period of two months, January and February. For consistency all values are presented in degrees Fahrenheit. It is clear that the HDD index exhibits large variability. Similar the difference between the maximum and minimum is close to 70 degrees Fahrenheit in average for all cities while the standard deviation of temperature is close to 15 degrees Fahrenheit. Also for all cities there is kurotsis significant smaller than 3 and with exceptions of Barcelona Madrid and London there is negative skewness.

Table 1. Descriptive Statistics of temperature in each city

	Mean	St.Dev	Max	Min	Skewness	Kurtosis	Mean HDD	std. HDD
Paris	54.38	12.10	89.90	13.80	-0.04	2.50	1368.40	134.95
Rome	60.20	11.37	85.80	31.10	-0.04	1.96	1075.00	121.52
Stockholm	45.51	14.96	79.20	-5.00	-0.09	2.33	2114.36	197.93
Amsterdam	51.00	11.00	79.90	12.20	-0.18	2.54	1512.07	189.05
Barcelona	61.56	10.59	85.70	32.60	0.09	2.03	899.23	102.75
Madrid	58.61	13.84	89.80	24.90	0.17	1.94	1262.56	128.53
New York	55.61	16.93	93.70	8.50	-0.15	2.08	1783.44	207.13
London	52.87	10.03	83.00	26.70	0.02	2.36	1307.11	98.80
Oslo	41.47	15.65	74.60	-8.70	-0.31	2.50	2404.53	236.19
Atlanta	62.18	14.52	89.60	13.70	-0.45	2.25	1130.75	121.19
Chicago	50.61	19.40	91.40	-12.90	-0.25	2.17	2221.33	211.89
Portland	46.80	17.36	83.20	-3.70	-0.22	2.22	2382.04	192.47
Philadelphia	56.02	17.13	90.50	9.50	-0.19	2.03	1766.98	211.65

[2] Atlanta, Detroit, New York, Baltimore, Houston, Philadelphia, Boston, Jacksonville, Portland, Chicago, Kansas City, Raleigh, Cincinnati, Las Vegas, Sacramento, Colorado Spring, Little Rock, Salt Lake City, Dallas, Los Angeles, Tucson, Des Moines, Minneapolis-St. Paul, Washington, D.C..

[3] Amsterdam, Barcelona, Berlin, Essen, London, Madrid, Paris, Rome, Stockholm, Oslo.

[4] Tokyo, Osaka.

[5] Calgary, Montreal, Vancouver, Edmonton, Toronto, Winnipeg.

Next we forecast the two months, 59 days, ahead out-of-sample forecasts for the CAT and cumulative HDD indices. Our method is validated and compared against two forecasting methods proposed in prior studies, the historical burn analysis (HBA) and the Benth's & Saltyte-Benth's (B-B) model which is the starting point for our methodology.

Table 2 shows the relative (percentage) errors for the CAT index of each method. It is clear that the proposed method using WN outperforms both HBA and B-B. More precisely the WN give smaller out-of-sample errors in 9 out of 13 times while it outperforms B-B in 11 out of 13 times. It is clear that the WN can be used with great success in European cities where the WN produces significant smaller errors than the alternative methods. Only in Oslo and Amsterdam WN performs worse than the HBA method but still the forecasts are better than the B-B. In USA cities WN produces the smallest out of sample error in three cases while HBA and B-B produce the smaller out of sample error in one and two cases respectively. Observing Table 1 again, we notice that when the temperature shows large negative skewness, with exception of New York, Portland and Philadelphia, the proposed method is outperformed either by HBA or by B-B. On the other hand in the cases of Barcelona and Madrid where the skewness is positive the errors using the wavelet network method are only 0.03% and 0.74% and significant smaller than the errors produced by the other two methods. Table 3 shows the relative (percentage) errors for the HDD index of each method. The results are similar.

Finally we examine the fitted residuals in model (6). Note that the B-B model, in contrast to the wavelet network model, is based on the hypothesis that the remaining residuals follow the normal distribution. It is clear from Table 4 that only in Paris the normality hypothesis marginally accepted. The Jarque-Bera statistic is slightly higher than 0.05. In every other case the normality test is rejected. More precisely the Jarque-Bera statistics are very large and the *p*-values are close to zero. Hence, alternative methods like wavelet analysis, must be used to capture the seasonal part of the data, [5].

Table 2. Relative errors for the three forecasting models. CAT index.

Errors	HBA	B-B	WNN
Paris	10.63%	8.34%	7.12%
Rome	4.49%	4.39%	3.93%
Stockholm	10.34%	9.47%	9.29%
Amsterdam	7.40%	8.60%	8.55%
Barcelona	1.46%	0.19%	0.03%
Madrid	7.18%	2.10%	0.74%
New York	10.63%	9.02%	8.76%
London	7.99%	6.07%	5.75%
Oslo	2.87%	5.62%	4.53%
Atlanta	2.21%	1.83%	2.58%
Chicago	15.55%	10.22%	10.90%
Portland	14.02%	8.87%	8.32%
Philadelphia	8.45%	5.95%	5.92%

Table 3. Relative errors for the three forecasting models. HDD index.

	HBA	B-B	WNN
Paris	14.77%	11.58%	9.88%
Rome	9.94%	9.71%	8.70%
Stockholm	7.09%	6.49%	6.37%
Amsterdam	9.57%	11.13%	11.06%
Barcelona	4.50%	0.57%	0.10%
Madrid	12.00%	3.51%	1.24%
New York	15.86%	13.45%	13.07%
London	12.52%	9.51%	9.01%
Oslo	1.64%	3.20%	2.58%
Atlanta	5.71%	4.74%	6.67%
Chicago	15.45%	10.15%	10.82%
Portland	11.05%	6.99%	6.56%
Philadelphia	12.11%	8.53%	8.49%

Table 4. Normality test for the B-B residuals

	Jarque-Bera	P-Value
Paris	5.7762	0.054958
Rome	170.12339	0.001
Stockholm	60.355	0.001
Amsterdam	44.6404	0.001
Barcelona	685.835	0.001
Madrid	69.52	0.001
New York	53.91428	0.001
London	11.66163	0.003947
Oslo	37.272738	0.001
Atlanta	403.0617	0.001
Chicago	44.329798	0.001
Portland	21.91905	0.001
Philadelphia	89.54923	0.001

5 Temperature Derivative Pricing.

So far, we modeled the temperature using an Ornstein-Uhlenbeck process [9]. We have shown in [5] that the mean reversion parameter α in model (6) is characterized by significant daily variation. Recall that parameter α is connected to our initial model with $a = e^{-\kappa}$ where κ is the speed of mean reversion. It follows that, the assumption of a constant mean reversion parameter introduces significant error in the pricing of weather derivatives. In this section, we give the pricing formula for a future contract written on the HDD index that incorporate the time dependency of the speed of the mean reversion parameter. The corresponding equations for the CAT index already presented in [5].

The CDD, HDD indices over a period $[\tau_1, \tau_2]$ are given by

$$HDD = \int_{\tau_1}^{\tau_2} \max\left(c - T(s), 0\right) ds \tag{16}$$

$$CDD = \int_{\tau_1}^{\tau_2} \max\left(T(s) - c, 0\right) ds \tag{17}$$

Hence, the pricing equations are similar for both indices.

First, we re-write (1) where parameter *κ*, now is a function of time *t*, *κ(t)*.

$$dT(t) = dS(t) + \kappa(t)\left(T(t) - S(t)\right) + \sigma(t)dB(t) \tag{18}$$

From the Ito formula an explicit solution can be derived:

$$T(t) = S(t) + e^{\int_0^t \kappa(u)du}\left(T(0) - S(0)\right) + e^{\int_0^t \kappa(u)du} \int_0^t \sigma(s) e^{-\int_0^s \kappa(u)du} dB(s) \tag{19}$$

Note that $\kappa(t)$ is bounded away from zero.

Our aim is to give a mathematical expression for the HDD future price. It is clear that the weather derivative market is an incomplete market. Cumulative average temperature contracts are written on a temperature index which is not a tradable or storable asset. In order to derive the pricing formula, first we must find a risk-neutral probability measure *Q~P*, where all assets are martingales after discounting. In the case of weather derivatives any equivalent measure *Q* is a risk neutral probability. If *Q* is the risk neutral probability and *r* is the constant compounding interest rate then the arbitrage free future price of a HDD contract at time $t \le \tau_1 \le \tau_2$ is given by:

$$e^{-r(\tau_2 - t)} \mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \max\left(0, c - T(\tau)\right) d\tau - F_{HDD}(t, \tau_1, \tau_2) \mid \mathfrak{F}_t\right] = 0 \tag{20}$$

and since F_{HDD} is $\mathfrak{F}_t$ adapted we derive the price of a HDD futures to be

$$F_{HDD}(t, \tau_1, \tau_2) = \mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \max\left(0, c - T(\tau)\right) d\tau \mid \mathfrak{F}_t\right] \tag{21}$$

Using the Girsanov's Theorem, under the equivalent measure *Q*, we have that

$$dW(t) = dB(t) - \theta(t)dt \tag{22}$$

and note that $\sigma(t)$ is bounded away from zero. Hence, by combining equations (18) and (22) the stochastic process of the temperature in the risk neutral probability *Q* is:

$$dT(t) = dS(t) + \left(\kappa(t)\left(T(t) - s(t)\right) + \sigma(t)\theta(t)\right)dt + \sigma(t)dW(t) \tag{23}$$

where $\theta(t)$ is a real-valued measurable and bounded function denoting the market price of risk. The market price of risk can be calculated by historical data. More specifically $\theta(t)$ can be calculated by looking the market price of contracts. The value that makes the price of the model fits the market price is the market price of risk. Using Ito formula, the solution of equation (23) is:

$$T(t) = S(t) + e^{\int_0^t \kappa(u)du}\big(T(0) - S(0)\big) + e^{\int_0^t \kappa(u)du}\int_0^t \sigma(s)\theta(s)e^{-\int_0^s \kappa(u)du}ds \\ + e^{\int_0^t \kappa(u)du}\int_0^t \sigma(s)e^{-\int_0^s \kappa(u)du}dB(s) \tag{24}$$

By replacing this expression to (21) we find the price of future contract on HDD index at time t where $0 \le t \le \tau_1 \le \tau_2$. Following the notation of [28] we have the following proposition.

Proposition 1. The HDD future price for $0 \le t \le \tau_1 \le \tau_2$ is given by

$$F_{HDD}(t,\tau_1,\tau_2) = \mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \max\big(c - T(s)\big)ds \,|\mathcal{F}_t\right] = \int_{\tau_1}^{\tau_2} v(t,s)\Psi\left(\frac{m(t,s)}{v(t,s)}\right)ds \tag{25}$$

where,

$$m(t,s) = c - S(s) - e^{\int_t^s \kappa(z)dz}\tilde{T}(t) - e^{\int_t^s \kappa(z)dz}\int_t^s \sigma(u)\theta(u)e^{-\int_t^u \kappa(z)dz}du \tag{26}$$

$$v^2(t,s,x) = e^{2\int_t^s \kappa(z)dz}\int_t^s \sigma^2(u)\theta(u)e^{-2\int_t^u \kappa(z)dz}du \tag{27}$$

and $\Psi(x) = x\Phi(x) + \Phi'(x)$ where Φ is the cumulative standard normal distribution function.

Proof. From equation (21) and (24) we have that:

$$F_{HDD}(t,\tau_1,\tau_2) = \mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \max\big(c - T(s)\big)ds \,|\mathcal{F}_t\right]$$

and using Ito's Isometry we can interchange the expectation and the integral

$$\mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \max\big(c - T(s)\big) |\mathcal{F}_t\right] == \int_{\tau_1}^{\tau_2} \mathrm{E}_Q\left[\max\big(c - T(s)\big) \,|\, \mathcal{F}_t\right]ds$$

$T(s)$ is normally distributed under the probability measure Q with mean and variance given by:

$$E_Q\left[T(s) \,|\, \mathcal{F}_t\right] = S(s) + e^{\int_t^s \kappa(z)dz}\tilde{T}(t) + e^{\int_t^s \kappa(z)dz}\int_t^s \sigma(u)\theta(u)e^{-\int_t^u \kappa(z)dz}du$$

$$Var_Q\left[T(s) \,|\, \mathcal{F}_t\right] = e^{2\int_t^s \kappa(z)dz}\int_t^s \sigma^2(u)\theta(u)e^{-2\int_t^u \kappa(z)dz}du$$

Hence, $c - T(s)$ is normally distributed with mean given by $m(t,s)$ and variance given by $v^2(t,s)$ and the proposition follows by standard calculations using the properties of the normal distribution.

6 Conclusions

This paper proposes and implements a modeling and forecasting framework for temperature based weather derivatives. The proposed method is an extension of the works proposed by [5] and [9]. Here the speed of mean reversion parameter is considered to be a time varying parameter and it is modeled by a wavelet neural network. It is proved that the waveform of the activation function of the proposed network provides a better fit of the data.

Our method is validated in a two month ahead out of sample forecast period. Moreover the relative errors produced by the wavelet network are compared against the original B-B model and historical burn analysis. Results show that the wavelet network outperforms the other methods. More precisely the wavelet network forecasting ability is better than the B-B and HBA in 11 times out of 13. Finally testing the fitted residuals of B-B we observe that the normality hypothesis is rejected in almost every case. Hence, B-B cannot be used for forecasts. Finally, we provided the pricing equations for temperature futures of a HDD index derivative when α is time depended.

The results in this study are preliminary and can be improved. More precisely the number of sinusoids in equations (2) and (3) in B-B framework, representing the seasonal part of the temperature and the variance of residuals, are chosen according to [9]. Alternative methods can improve the fitting in the original data. Hence a better training set is expected for the wavelet network and more accurate forecasts.

Another important aspect is to test the largest forecasting window of each method. Meteorological forecasts of a window larger than 10 days considered inaccurate. Hence, it is important to develop a model than can accurately predict daily average temperatures in larger windows. Also, this analysis will let us use the best model according to the desired forecasting interval.

References

1. Challis, S.: Bright Forecast for Profits. Reactions **June edition** (1999)
2. Hanley, M.: Hedging the Force of Nature. Risk Professional **1** (1999) 21-25
3. Ceniceros, R.: Weather derivatives running hot. Business Insurance **40** (2006)
4. Jewson, S., Brix, A., Ziehmann, C.: Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations. Campbridge University Press, Cambridge, UK (2005)
5. Zapranis, A., Alexandridis, A.: Modelling Temperature Time Dependent Speed of Mean Reversion in the Context of Weather Derivetive Pricing. Applied Mathematical Finance, Vol. 15 (2008) 355 - 386

6. Zapranis, A., Alexandridis, A.: Weather Derivatives Pricing: Modelling the Seasonal Residuals Variance of an Ornstein-Uhlenbeck Temperature Process With Neural Networks. Neurocomputing (accepted, to appear)
7. Alaton, P., Djehince, B., Stillberg, D.: On Modelling and Pricing Weather Derivatives. Applied Mathematical Finance **9** (2000) 1-20
8. Zhang, Q., Benveniste, A.: Wavelet Networks. IEEE Trans. Neural Networks **3** (1992) 889-898
9. Benth, F.E., Saltyte-Benth, J.: The volatility of temperature and pricing of weather derivatives. Quantitative Finance **7** (2007) 553-561
10. Daubechies, I.: Ten Lectures on Wavelets. SIAM, Philadelphia, USA (1992)
11. Mallat, S.G.: A Wavelet Tour of Signal Processing. Academic Press, San Diego (1999)
12. Zapranis, A., Alexandridis, A.: Wavelet analysis and weather derivatives pricing. HFFA, Thessaloniki (2006)
13. Oussar, Y., Dreyfus, G.: Initialization by Selection for Wavelet Network Training. Neurocomputing **34** (2000) 131-143
14. Zapranis, A., Alexandridis, A.: Model Identification in Wavelet Neural Networks Framework. In: Iliadis, L., Vlahavas, I., Bramer, M. (eds.): Artificial Intelligence Applications and Innovations III, Vol. IFIP 296. Springer, New York, USA (2009) 267-277
15. Cao, M., Wei, J.: Pricing the weather. Risk Weather Risk Special Report, Energy And Power Risk Management (2000) 67-70
16. Davis, M.: Pricing weather derivatives by marginal value. Quantitative Finance **1** (2001) 1-4
17. Dornier, F., Queruel, M.: Caution to the wind. Weather risk special report, Energy Power Risk Management (2000) 30-32
18. Moreno, M.: Riding the temp. Weather Derivatives, FOW Special Support (2000)
19. Caballero, R., Jewson, S., Brix, A.: Long Memory in Surface Air Temperature: Detection Modelling and Application to Weather Derivative Valuation. Climate Research **21** (2002) 127-140
20. Brody, C.D., Syroka, J., Zervos, M.: Dynamical Pricing of Weather Derivatives. Quantitave Finance **2** (2002) 189-198
21. Benth, F.E., Saltyte-Benth, J.: Stochastic Modelling of Temperature Variations With a View Towards Weather Derivatives. Applied Mathematical Finance **12** (2005) 53-85
22. Oussar, Y., Rivals, I., Presonnaz, L., Dreyfus, G.: Trainning Wavelet Networks for Nonlinear Dynamic Input Output Modelling. Neurocomputing **20** (1998) 173-188
23. Zhang, Q.: Using Wavelet Network in Nonparametric Estimation. IEEE Trans. Neural Networks **8** (1997) 227-236
24. Postalcioglu, S., Becerikli, Y.: Wavelet Networks for Nonlinear System Modelling. Neural Computing & Applications **16** (2007) 434-441
25. Xu, J., Ho, D.W.C.: A Basis Selection Algorithm for Wavelet Neural Networks. Neurocomputing **48** (2002) 681-689
26. Gao, R., Tsoukalas, H.I.: Neural-wavelet Methodology for Load Forecasting. Journal of Intelligent & Robotic Systems **31** (2001) 149-157
27. Xu, J., Ho, D.W.C.: A Constructive Algorithm for Wavelet Neural Networks. Lecture Notes in Computer Science (2005) 730-739
28. Benth, F.E., Saltyte-Benth, J., Koekebakker, S.: Putting a price on temperature. Scandinavian Journal of Statistics, **34** (2007) 746-767