Modeling Temperature Time-Dependent Mean Reversion with Neural Networks in the Context of Derivatives Pricing

A. Zapranis, A. Alexandridis

Abstract—In this paper, in the context of an Ornstein-Uhlenbeck temperature process we use neural networks to examine the time dependence of the speed of the mean reversion parameter α of the process. We estimate non-parametrically with a neural network a model of the temperature process and then we compute the derivative of the network output w.r.t. the network input, in order to obtain a series of daily values for α. To our knowledge, this is done for the first time, and it gives us a much better insight in temperature dynamics and in temperature derivative pricing. Our results indicate strong time dependence in the daily values of α but no seasonal patterns. This is important, since in all relevant studies so far, α was assumed to be constant. Furthermore, the residuals of the neural network provide a better fit to the normal distribution, when compared with the residuals of the classic linear models which are being used in the context of temperature modeling (where α is constant). It follows, that by setting the mean reversion parameter to be a function of time we improve the accuracy of the pricing of the temperature derivatives. Finally, we provide the pricing equations for temperature futures and options, when α is time dependent.

Index Terms—Neural networks, Weather derivatives pricing.

1 Introduction

Temperature derivatives have as an underlying variable, temperature indices such as Heating Degree Days (HDD) or Cooling Degree Days (CDD) defined on average daily temperatures. The list of traded contracts is extensive and constantly evolving. In the Chicago Mercantile Exchange (CME) there are traded weather contracts based on an index of Cumulative Average Temperature (CAT) for European cities for May to September. A CAT index is defined as the sum of the daily average temperatures over the period of the contract.

However, pricing weather derivatives is far from a straightforward task, since the underlying weather index cannot be traded. Furthermore, the corresponding market is relatively illiquid. The weather derivatives market is a classic incomplete market, meaning that prices cannot be derived from the no-arbitrage condition, since it is not possible to replicate the payoff of a given contingent claim by a controlled portfolio of the basic securities.

In pricing a weather derivative, dynamic modeling of the daily temperatures is generally considered more appropriate than modeling the temperature index. In principle, it leads to more accurate pricing, but 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 descretized in order to estimate its various parameters. Once the process is estimated, one can then value any contingent claim by taking expectation of the discounted future payoff.

In this paper, we extent the mean-reverting process with seasonality in the level and volatility, proposed by Benth and Saltyte-Benth [2] - a generalization of the process proposed earlier by Dornier and Querel [10], which is descretized in the form of an AR(1) model. We estimate non-parametrically a non-linear AR(1) model with a neural network. This removes the constraint of a constant mean reverting parameter. By computing the derivative of the network output w.r.t. the network input, we take a series of daily values for the mean reversion parameter for a period of 30 years for the city of Paris. Analytical expressions for the various network derivatives are given by Zapranis and Refenes [17].

It is important to mention here, that up to date the mean reversion parameter was assumed constant in all relevant studies. However, our findings indicate exactly the opposite. The daily variation of the value of the mean reversion parameter is quite high. The non-linear neural model which encapsulated this time dependency provides a much better fit to the temperature data than the classic linear alternative. The implications in the accuracy of the pricing process of this type of derivatives are obvious. Furthermore, the complexity of

- *A. Zapranis is with the Department of Accounting and Finance, University of Macedonia of Economics and Social Studies, 156 Egnatia St., 54006, Thessaloniki, Greece. E-mail: zapranis@uom.gr.*
- *A. Alexandridis is with the Department of Accounting and Finance, University of Macedonia of Economics and Social Studies, 156 Egnatia St., 54006, Thessaloniki, Greece. E-mail: aalex@uom.gr.*

Manuscript received (11.6.07).

the pricing equations is not being increased significantly by using a time dependent mean reversion parameter. Below, first we describe the basic steps of our analysis and then the organization of the rest of the paper.

Given the temperature model, the first step is to identify and remove from the temperature series the (possible) trend and the non-stationary seasonal cycle, hoping that what is left will be stationary. This is usually done by modeling the seasonal variations as deterministic and the same every year (seasonally stationary). The stochastic variability of the temperature is then moved entirely from the seasonal cycle into the residuals.

In modeling the seasonal cycle deterministically, there are several approaches. The discrete Fourier transform (DTF) is considered to be the most accurate, since, in principle at least, removes the seasonal cycle both in the mean and in the variance. For a detailed discussion on this subject see Jewson and Brix [11]. However, recently Zapranis and Alexandridis [15],[16] proposed a novel approach in modeling the seasonal cycle which is an extension of the DFT approach. More specifically, we use wavelet analysis (WT), which superimposes sines and cosines to represent other functions, to decompose the temperature series into a series of (orthogonal) basis functions (wavelets) with different time and frequency locations. As a result, the wavelet decomposition brings out the structure of the underlying temperature series as well as trends, periodicities, singularities or jumps that could not be observed originally [1], [8]. Our approach was tested in 40 years of temperature data collected from Paris (from 1961 to 2000), and the improvement in terms of distributional properties was found to be significant. The same approach is used in this paper.

Once the trend and the seasonal cycle in the mean and the variance have being removed, one has to investigate the distributional properties of the residuals (anomalies) of the temperature process. To the extent that this part of the modeling approach and the initial temperature process are accurate, the residuals must follow a normal distribution with mean zero and standard deviation of one at all times of the year. However, often the hypothesis of normality is rejected [3].

As it is shown in the next section, the temperature process can be written as an AR(1) model after removing the linear trend and the seasonal component. Or, as we propose here as a non-linear AR(1) fitted non-parametrically with a neural network, which allows us to examine the time structure of the speed of the mean reversion of the temperature process. We show that temperature is a mean reverting process where the speed of mean reversion depends on time. Our findings were compared against a linear AR(1) process with a constant parameter.

Since, there is time dependency in the variance of the residuals we have to extract that variance. In doing so, we group the residuals in 365 groups, each group corresponding to a particular day of the year. Each group comprises 30 observations. Each observation corresponds to a different year. Then we take the average for each group. Using those 365 values we model the residual variance with a neural network having as inputs the harmonics corresponding to the seasonal cycles of the residuals, identified by a second wavelet analysis.

The rest of the paper is organized as follows. In section 2, we describe the process used to model the average daily temperature. The calibration of the temperature model is done based on the results of a wavelet analysis. We also estimate and then remove the linear trend and the seasonality component. In section 3, we model the speed of the mean reversion of the temperature process. In section 3.1 we use neural networks to extract daily values of the speed of mean reversion of our process. Then, we model the seasonal residual variance, again using the wavelet analysis approach. The analysis is repeated in section 3.2 a linear AR(1) model with a constant speed of mean reversion parameter. In section 4, we give the analytic expressions for pricing temperature futures and options with a time dependent reversion parameter. Finally, in section 5 we conclude.

2 Dynamic Modeling Of The 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 (see for example [1], [7], [8]). Other researchers (e.g., [10], [13]) have suggested versions of a more general ARMA(p,q) model. However, it has been shown 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 [6]. In order to deal with this problem, more complex models were proposed, with a characteristic example being the model of Brody *et al* [5], which is an Ornstein-Uhlenbeck process. This model was further extended, at first by replacing the noise part of the process (Brownian) by a fractional Brownian noise and then by a Levy process [3].

Our analysis is based on the model of Benth and Saltyte-Benth, where the temperature is expressed as a mean reverting Ornstein-Uhlenbeck process, i.e.

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

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

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

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

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 $S(t)$. In this paper, the exact specification of models (2) and (3) is decided based on the results of wavelet analysis of the temperature series.

In this section we derive the characteristics and dynamics of the daily temperature of the city of Paris, France. The data consists average daily temperatures of 30 years (1971-2000). The distribution of the data is not normal, indicating a temperature process that is generally hard to model.

In order to identify the number of terms I_1, J_1 in (2) and I_2, J_2 in (3) we decompose the temperature series using a wavelet transform (WT), a generalization of the DFT and the windowed Fourier (WFT) transform. The wavelet transform is localized in both time and frequency. Also it adapts itself to capture features across a wide range of frequencies, thus avoiding the assumption of stationarity. In addition, wavelets have the ability to decompose a signal or a time-series in different levels.

At each level j, we build the j-level approximation a_j, or *approximation* at level j, and a deviation signal called the j-level detail d_j, or *detail* at level j. We can consider the original signal as the approximation at level 0, denoted by a_0. The words approximation and detail are justified by the fact that a_1 is an approximation of a_0 taking into account the low frequencies of a_0, whereas the detail d_1 corresponds to the high frequency correction. For detailed expositions on the mathematical aspects of wavelets we refer to [9], [12], [14].

For the decomposition of the average daily temperature time-series the Daubechies 11 wavelet at level 11 was used. As shown in these papers an upward trend exists in the temperature in Paris. Also the a series of cycles affects the dynamics of temperature. An one year cycle exists, as expected. Moreover, cycles of 2, 4, 8 and 13 years also exists and affect the temperature dynamics. Also, wavelet analysis captures a product of two sinusoids, with a period of 1 and 7 years respectively. Finally, the lower details reflect the noise part of the time-series. A closer inspection of the noise part reveals seasonalities, which will be extracted later on.

A discrete approximation to (4), which is the solution to the mean reverting Ornstein-Uhlenbeck process (1), is:

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

which can be written as:

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

where

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

$$\tilde{\sigma}(t) = a\sigma(t) \quad (8)$$

$$a = e^{-k} \quad (9)$$

In order to estimate (6) we need first to remove the trend and seasonality components from the average temperature series.

Firstly, we quantify the upward trend indicated by the results of the wavelet analysis by fitting a linear regression to the temperature data. The regression is statistically significant with intercept 0.00010562 and slope 11.171. Subtracting the trend form the original data we obtain the de-trended temperature series.

The results of the wavelet analysis indicate that the seasonal part of the temperature takes the following form:

$$\begin{aligned} S(t) = a &+ b_1 \sin(2\pi(t - f_1)/365) \\ &+ b_2 \sin(2\pi(t - f_2)/(2 \cdot 365)) \\ &+ b_3 \sin(2\pi(t - f_3)/(13 \cdot 365)) \\ &+ b_4 (1 + \sin(2\pi(t - f_4) \\ &/(7 \cdot 365))) \sin(2\pi t/365) \\ &+ b_5 \sin(2\pi(t - f_5)/(8 \cdot 365)) \\ &+ b_6 \sin(2\pi(t - f_6/(4 \cdot 365)) \end{aligned} \quad (10)$$

The estimated parameters of the above model are as follows: a = -0.0008, b_1 = -7.6994, b_2 = 0.1317, b_3 = 0.0469, b_4 = -0.2743, b_5 = -0.3445, b_6 = 0.0796, f_1 = -73.2644, f_2 = 95.0642, f_3 = -640.2319, f_4 = 183.1090, f_5 = -13.1151 and f_6 = -134.5803. The mean of the residuals 5.9091e-009 and the standard deviation is 3.3708. Next the temperature series is de-seasonalized by removing $S(t)$.

3 Modeling the Mean Reverting Parameter

The de-trended and de-seasonalized temperature series can be modeled with an AR(1) process with a zero constant term, as shown in (6). In the context of such a model the mean reversion parameter a is typically assumed to be constant over time. In [5] it is mentioned that in general a should be a function of time, but no evidence was presented. On the other hand, Benth and Saltyte-Benth [3], using a dataset comprising 10 years of Norwegian temperature data, calculated mean annual values of α. They reported that their variation from year to year was not significant. They also investigated the seasonal structures in monthly averages of α and they reported that none was found. However, since to date, no one has computed daily values of the mean reversion parameter, since there is no obvious way to do this in the context of model (6). On the other hand, averaging techniques, in a yearly or monthly basis, run the danger of filtering out too much variation and consequently presenting a distorted picture regarding the true nature of a. The impact of a false specification of a, on the accuracy of the pricing of temperature derivatives is significant [1].

In this section, we address that issue, by using a neural network to estimate non-parametrically relationship (6) and then estimate a as a function of time. By computing the derivative of the network output w.r.t. the network input we obtain a series of daily values for a. This is done for the first time, and it gives us a much better insight in temperature dynamics and in temperature derivative pricing. As we will see the daily variation of a is quite significant after all.

3.1 The Neural Networks Approach: Time Dependent Mean Reversion Parameter

Our temperature data consisted of 30 years up to 2000, of de-trended and de-seasonalized daily average temperatures from the city of Paris. We separated the data in 3 groups, each group corresponding to one decade. Then, using neural networks we estimated non-parametrically the generalized version of (6), that is:

$$T(t+1) = \varphi\big(T(t)\big) + e(t) \tag{11}$$

Once we have the estimator $\hat{\varphi}$ of the underlying function φ, then we can compute the daily values of a as follows:

$$a(t) \equiv dT(t+1)/dT(t) = d\varphi/dT \tag{12}$$

The analytic expression for the neural network derivative $d\varphi/dT$ can be found in Zapranis & Refenes [17].

For the 3rd (last) decade the daily values of a (3,650 values) are depicted in Fig. 1. The corresponding frequency histogram is given in Fig. 2. The graphs for the 1st and 2nd decades are very similar. The relevant statistics for all three decades are given in Table 1.

It is clear, that the mean reversion parameter is not constant. On the contrary, its daily variation is quite significant; this fact naturally has an impact on the accuracy of the pricing equations and it has to be taken into account.

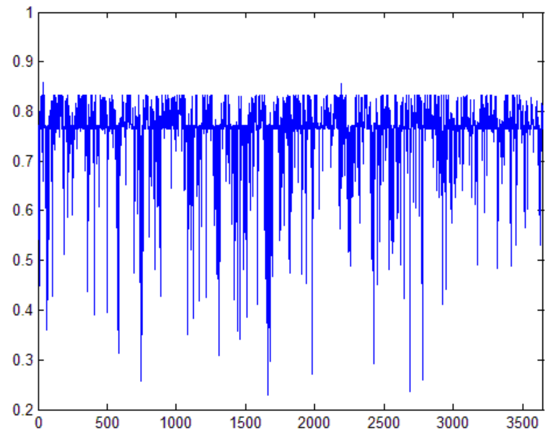

Fig.1. Daily variation of the mean reversion parameter α.

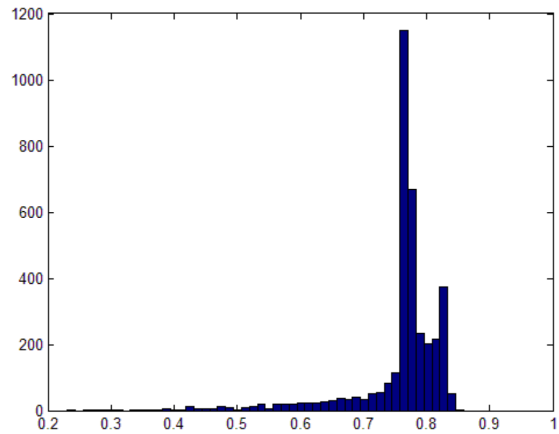

Fig.2. Frequency distribution of the mean reversion parameter α.

Referring now to Fig. 1, we observe that the spread between the maximum and minimum value is quite high (0.8586 and 0.2303 correspondingly). The standard deviation is 0.0587 and the mean is 0.7573. We also observe, that there is an upper threshold in the values of a (0.8376) which is rarely exceeded. This can also be seen in the frequency distribution of a in Fig. 2. A closer examination of a did not reveal any seasonalities.

The distributional statistics of the residuals of the neural network (Fig. 3), do not indicate a significant deviation from the normal distribution. There is a small negative skewness (-0.094027), positive kurtosis (3.031307) and the value of the Jarque-Bera statistic is 5.525856. The probability is 0.063107 (>0.05), indicating that we have to accept the normality hypothesis. The autocorrelation of the residuals is significant in the first lag (Fig. 4), while the autocorrelation of the squared residuals indicates time dependency in the variance of the residuals (Fig. 5). In Fig. 5, we can clearly observe the seasonal variation.

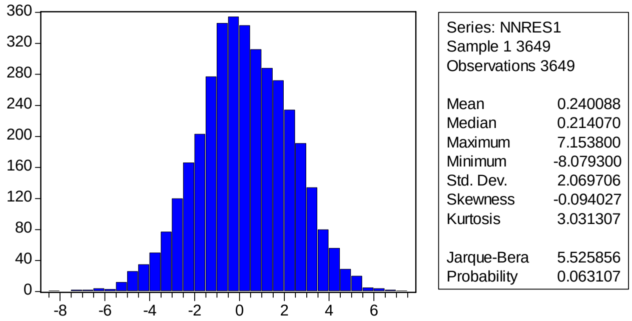

Fig.3. Distribution statistics of the residuals of the N.N. model

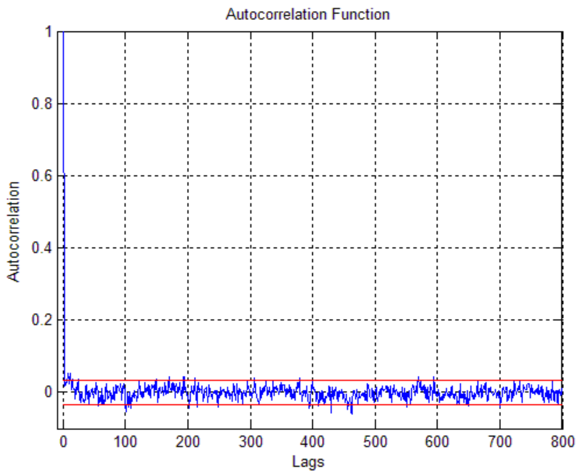

Fig.4. ACF of the residuals of the N.N. model for the de-trended and de-seasonalized Paris average daily data

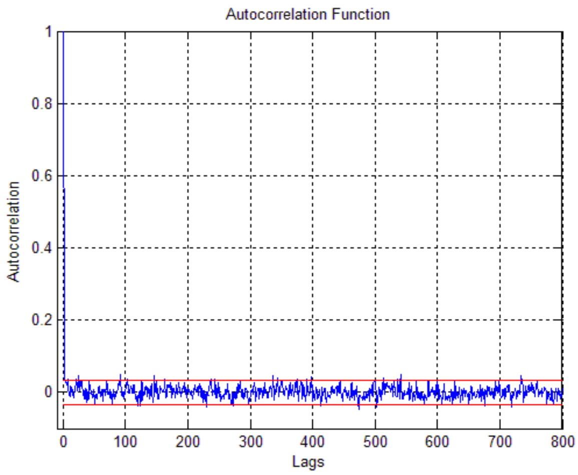

Fig.5. ACF of the squared residuals of the N.N. model for the de-trended and de-seasonalized Paris average daily data

Since, for the residuals $e(t)$ of the neural network model it is true that

$$e(t) = \tilde{\sigma}^2(t)\varepsilon(t)$$

where $\varepsilon(t)$ are i.i.d. N(0,1), we can extract the variance $\tilde{\sigma}^2(t)$ as follows: Firstly, we group the residuals in 365 groups, comprising 10 observations each (each group corresponds to a single day of the year). Then, by taking the average of the 10 squared values we obtain the variance for that day. That is, we assume that

$$\sigma^2(365+t) = \sigma^2(t) \tag{13}$$

We also know, that

$$\sigma^2(t) = \tilde{\sigma}^2(t)/a(t)^2$$

where $t = 1, \ldots, 3{,}650$ (for each decade).

In deciding which terms of a truncated Fourier series to use in order to model the variance $\sigma^2(t)$, we perform again a wavelet analysis, which indicates the presence of five cycles within $\sigma^2(t)$. The wavelet decomposition of the seasonal variance is shown in Fig. 6. Approximation a_7 and details d_7, d_6, d_5 suggest an one-year cycle, a half-year cycle, a 1/4 of a year cycle, a 1/9 of a year cycle and a 1/18 of a year cycle, respectively. We model accordingly the variance $\sigma^2(t)$, as follows:

$$\begin{aligned}\sigma^2(t) = c_0 &+ c_1\sin(2\pi t/365)\\ &+c_2\sin(4\pi t/365)\\ &+c_3\sin(8\pi t/365)\\ &+c_4\sin(18\pi t/365)\\ &+c_5\sin(36\pi t/365)\\ &+d_1\cos(2\pi t/365)\\ &+d_2\cos(4\pi t/365)\\ &+d_3\cos(8\pi t/365)\\ &+d_4\cos(18\pi t/365)\\ &+d_5\cos(36\pi t/365)\end{aligned} \tag{14}$$

The values of the estimated parameters of (14) for the 3rd decade are: c_0 = 4. 3390, c_1 = 0.5095, c_2 = -0.0721, c_3 = 0.1883, c_4 = 0.1533, c_5 =0.1379, d_1 = 0.1260, d_2 = 0.6230, d_3 = -0.2897 d_4 = 0.0637 and d_5 = -0.0431.

The empirical values of the variance of the residuals (365 values) together with the fitted variance

$$\tilde{\sigma}^2(t) = a^2(t)\sigma^2(t) \tag{15}$$

can be seen in Fig. 7. We observe that the variance takes its highest values during the winter months, while it takes its lowest values during early Autumn.

The standard deviation of the residuals is 2.0697, while the standard deviation of the remaining noise part is 0.9962 and its mean is 0.1165.

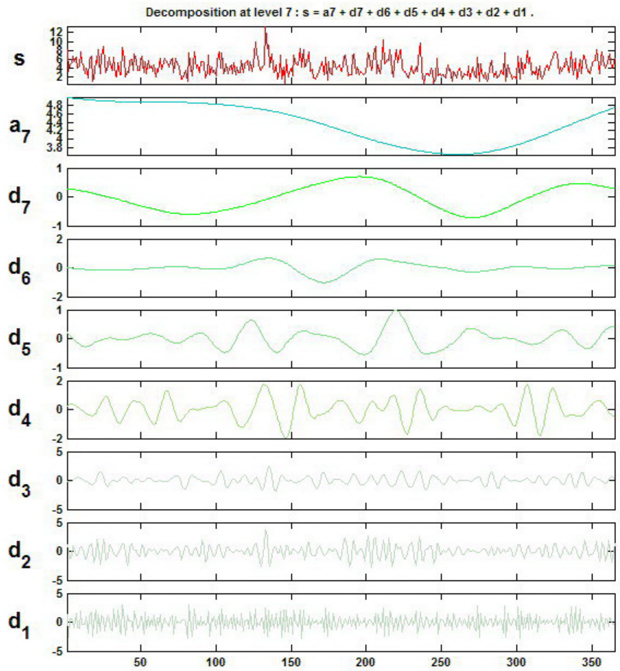

Fig.6. Wavelet decomposition of the averaged squared variance.

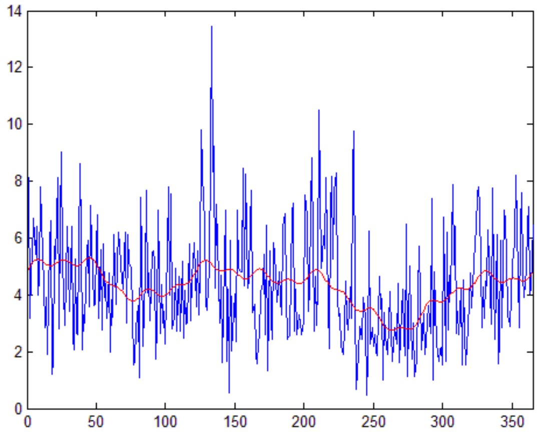

Fig.7. Empirical variance and fitted variance for the NN model.

In Fig. 8, we can see the autocorrelation function of the squared residuals of the process, after dividing out the volatility (15) from the residuals.

We observe that the seasonality has been removed, but there is still autocorrelation in the first lag. Moreover, the Jarque-Bera statistic is reduced to 2.568741 with a p-value of 0.276825 leading to the acceptance of the hypothesis of normal distribution.

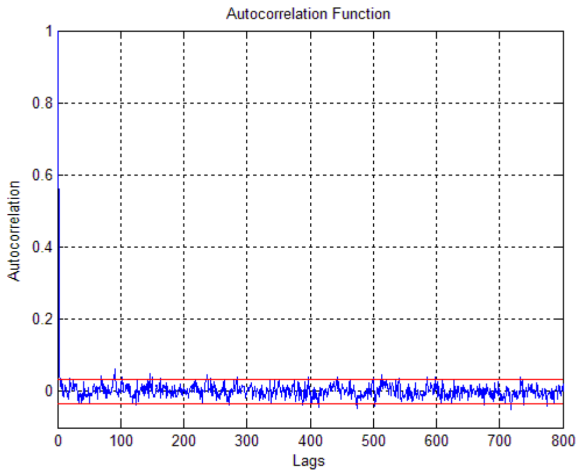

Fig.8. ACF of the squared residuals of the NN model after dividing out the volatility function from the residuals.

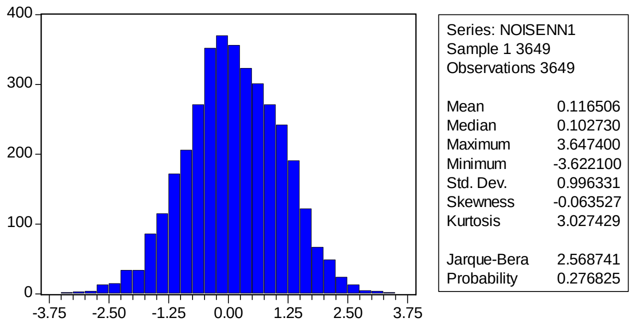

Fig.9. Distribution statistics of the residuals of the NN model after dividing out the volatility function from the residuals.

3.3 The Linear Regression Approach: Constant Mean Reversion Parameter

In the this section, we repeat the previous analysis but instead of a neural network we use the linear AR(1) model, proposed by Benth & Saltyte-Benth [2]. First we estimate the parameter α for the AR(1) model for each decade. For all three decades the constant was found to be zero, as it was expected, while the reversion parameter a takes the values: α_1 = 0.797, α_2 = 0.7989, α_3 = 0.8005 (the subscript indicates the decade); these values are also statistically significant (t = 79.35, 79.88, 80.33). As expected, the values of α are actually very close to the average values of $\alpha(t)$ which were derived from the neural network models. For all three decades the adjusted R^2 is over 0.63 and F is over 6297.

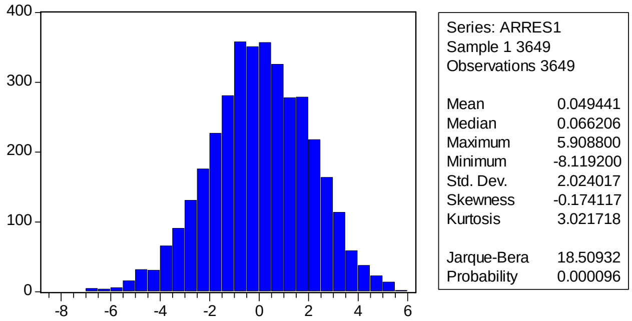

Fig.10. Distribution statistics of the residuals of the AR(1) model

The distributional statistics of the residuals of the AR(1) model (6), indicate a significant deviation from the normal distribution. There is a negative skewness (-0.174117), positive kurtosis (3.021718) and the value of the Jarque-Bera statistic is 18.50932. The *p*-value is less than 0.05, so that the hypothesis of normal distribution has to be rejected. It is clear, that the residuals obtained from the neural models are much closer to the normal distribution than the ones obtained from the AR(1) models.

Next, we estimate the seasonal variance. Now the value of a is obtained by a linear regression. For the last decade α_3 = 0.8005. The standard deviation of the remaining noise part is 1.0007 and the mean is 0.0253. The Jarque-Bera is 14.22022 and its *p*-value is 0.000817; again we have to reject the hypothesis of normal distribution.

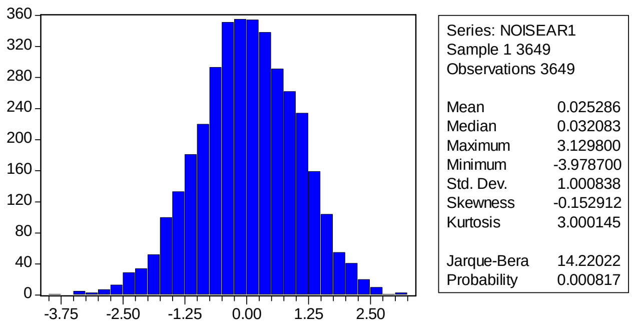

Fig.11. Distribution statistics of the residuals of the AR(1) model after dividing out the volatility function from the regression residuals

As we have seen the hypothesis of normality was accepted only in the case of the neural models. In Fig. 12 and Fig. 13 we can see the normality plots for the residuals (after dividing out the seasonal variance) of the neural network and the AR(1) model for the 3rd decade. Clearly, in the case of the neural model (Fig. 12) the residuals provide a better fit to the normal distribution.

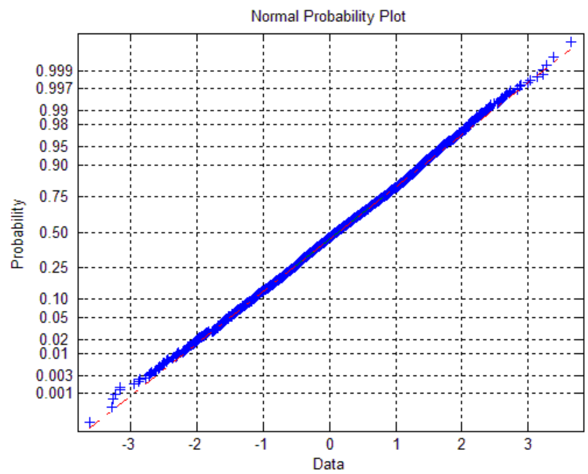

Fig.12. Normal probability plot of the of the residuals of the NN model after dividing out the volatility function from the residuals

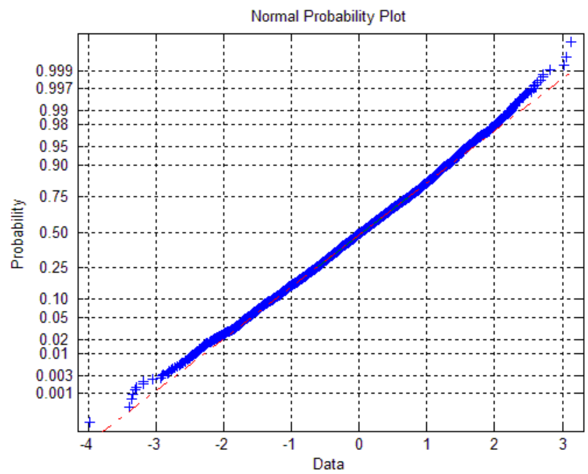

Fig.13. Normal probability plot of the of the residuals of the AR(1) model after dividing out the volatility function from the residuals

In Table 1, we can see the distributional statistics for all three decades. The neural networks approach always gives a smaller Jarque-Bera and higher *p*-value. Moreover, the skewness of the distributions corresponding to the NNs is always significantly lower, although, the kurtosis is lower only for the first decade. Finally, in the last decade the normality hypothesis using the linear regression is rejected (*p*=0.000817) while it is accepted using the neural network approach (*p*=0.276825).

4 Temperature Derivatives Pricing

So far, we modeled the temperature using an Ornstein-Uhlenbeck process (as in Benth and Saltyte-Benth [2]) and we also used wavelet analysis to identify and filter out the seasonal component. Moreover, we have shown that the mean reversion parameter a is characterized by significant daily variation. 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 formulae that incorporate the time dependency of the mean reversion parameter

The value of a CAT index for the time interval $[\tau_1,\tau_2]$ is given by the following expression:

$$\int_{\tau_1}^{\tau_2} \mathrm{T}(\tau)d\tau \tag{16}$$

where the temperature is measured in degrees of Celsius.

If Q is the risk neutral probability and r is the constant compounding interest rate then the future price of a CAT contract at time t will be:

$$e^{-r(\tau_2-t)}\mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \mathrm{T}(\tau)d\tau - F_{CAT}(t,\tau_1,\tau_2) \mid F_t\right]=0 \tag{17}$$

and

$$F_{CAT}(t,\tau_1,\tau_2)=\mathrm{E}_Q\left[\int_{\tau_1}^{\tau_2} \mathrm{T}(\tau)d\tau \mid F_t\right] \tag{18}$$

The stochastic process for the temperature is:

$$dT(t)=ds(t)+\big(\theta(t)-\kappa\big(T(t)-s(t)\big)\big)dt+\sigma(t)dW(t) \tag{19}$$

where $\theta(t)$ is a real-valued measurable and bounded function. The solution to this equation is:

$$\begin{aligned} T(t) &= s(t)+\big(T(0)-s(0)\big)e^{-\kappa t} \\ &+\int_0^t \theta(u)e^{-\kappa(t-u)}du \\ &+\int_0^t \sigma(u)e^{-\kappa(t-u)}dW(u) \end{aligned} \tag{20}$$

By replacing this expression to (16) we get:

$$\begin{aligned} &\int_{\tau_1}^{\tau_2}\mathrm{T}(\tau)d\tau=\int_{\tau_1}^{\tau_2} s(t)dt \\ &-\kappa^{-1}\big(T(0)-s(0)\big)\big(e^{-\kappa\tau_2}-e^{-\kappa\tau_1}\big) \\ &-\int_0^{\tau_2}\theta(t)\kappa^{-1}\left\{e^{-\kappa(\tau_2-t)}-1_{[0,\tau_1]}(t)e^{-\kappa(\tau_1-t)}-1_{[\tau_1,\tau_2]}(t)\right\}dt \\ &-\int_0^{\tau_2}\sigma(t)\kappa^{-1}\left\{e^{-\kappa(\tau_2-t)}-1_{[0,\tau_1]}(t)e^{-\kappa(\tau_1-t)}-1_{[\tau_1,\tau_2]}(t)\right\}dW(t) \end{aligned} \tag{21}$$

The future price of a CAT contract $F_{CAT}(t,\tau_1,\tau_2)$ at time $t \le \tau_1$ then is:

$$\begin{aligned} \Theta(t,\tau_1,\tau_2) &= \kappa^{-1}\int_t^{\tau_2}\theta(u)\big(1-e^{\kappa(\tau_2-u)}\big)du \\ &-\kappa^{-1}\int_t^{\tau_1}\theta(u)\big(1-e^{\kappa(\tau_1-u)}\big)du \end{aligned} \tag{23}$$

Similarly, we can derive the price of a call option. Since $F_{CAT}(t,\tau_1,\tau_2)$ is an additive Gaussian process, we can compute the price of a call option at time t that expires at τ and has strike price K. For $t \le \tau \le \tau_1$ is:

$$\begin{aligned} F_{CAT}(t,\tau_1,\tau_2) &= \int_{\tau_1}^{\tau_2} s(t)dt \\ &-\kappa^{-1}\big(T(t)-s(t)\big)\big(e^{-\kappa(\tau_2-t)}-e^{-\kappa(\tau_1-t)}\big) \\ &+\Theta(t,\tau_1,\tau_2) \end{aligned} \tag{22}$$

where $\Theta(t,\tau_1,\tau_2)$ is given by the expression:

$$C(t)=e^{-r(\tau-t)}(F_{CAT}(t,\tau_1,\tau_2)-\mathrm{K})\Phi(d)+\frac{\Sigma_{t,\tau}}{\sqrt{2\pi}}e^{-d^2/2} \tag{24}$$

where:

$$d=\frac{F_{CAT}(t,\tau_1,\tau_2)-K}{\Sigma_{t,\tau}}$$

$$\Sigma_{t,\tau}^2 := \int_t^{\tau}\Sigma^2(u,\tau_1,\tau_2)du$$

$$\Sigma(t,\tau_1,\tau_2) := \kappa^{-1}(e^{-\kappa(\tau_1-t)}-e^{-\kappa(\tau_2-t)})\sigma(t)$$

In the above pricing equations the parameter κ is time dependent, that is:

$$\begin{aligned} k = k(t) &= -\ln\big(a(t)\big) = \\ &= -\ln\big(\partial\tilde{\mathrm{T}}(t+1)/\partial\tilde{\mathrm{T}}(t)\big) \end{aligned} \tag{25}$$

TABLE 1
DISTRIBUTIONAL STATISTICS

	Decade1	Decade2	Decade3
Mean	-0.02237	-0.00374	0.025286
	0.027346	0.067015	0.116506
Median		-0.00012	0.032083
	0.039475	0.065229	0.10273
Maximum	3.6872	3.4283	3.1298
	3.6395	3.3923	3.6474
Minimum	-4.0096	-3.4861	-3.9787
	-3.7338	-3.3591	-3.6221
Std. Dev	1.000689	1.000795	1.000838
	1.000445	0.998902	0.996331
Skewness	-0.07411	-0.13531	-0.15291
	-0.05289	-0.11002	-0.06353
Kurtosis	3.10874	2.984192	3.000145
	3.05332	2.916664	3.027429
Jarque-Bera	5.139474	11.1759	14.22022
	2.134051	8.420029	2.568741
Probability	0.076556	0.003743	0.000817
	0.34403	0.014846	0.276825

Distributional statistic for each decade after dividing out the seasonal variance. The first row of each statistic corresponds to the AR(1) model. The second row corresponds to the N.N. model.

The derivative $d\widetilde{\mathrm{T}}(t+1)/d\tilde{T}(t)$ is computed analytically from the neural model equations. For the 3rd decade it is depicted in Fig. 1.

5 Conclusions

In this paper, in the context of an Ornstein-Uhlenbeck temperature process we have used neural networks to examine the time dependence of the speed of the mean reversion parameter α of the process. By computing the derivative $d\widetilde{\mathrm{T}}(t+1)/d\tilde{T}(t)$ of the fitted neural model, we obtained daily values for α. To our knowledge, we are the first to have done so. Our results, indicate strong time dependence in the daily values of α but no seasonal patterns. We compared the fit of the residuals to the normal distribution of two types of models. Neural networks, were a is a function of time, and AR(1) models, were a is constant. Generally, in the case of neural networks we have a better fit. It follows, that by setting the mean reversion parameter to be a function of time we improve the accuracy of the pricing of the temperature derivatives. Finally, we provided the pricing equations for temperature futures and options, when a is time dependent.

References

[1] P. Alaton, B. Djehince & D. Stillberg "On Modeling and Pricing Weather Derivatives", Applied Mathematical Finance, 9, pp 1-20, 2000.

[2] F.E. Benth & J. Saltyte-Benth "The Volatility Of Temperature And Pricing Of Weather Derivatives" *Pure Mathematics,* E-Print, 12, Department of Mathematics. University of Oslo, 2005.

[3] F.E. Benth & J. Saltyte-Benth "Stochastic Modeling of Temperature Variations With a View Towards Weather Derivatives", *Applied Mathematical Finance*, 12 (1), pp 53-85, 2005.

[4] A. Brix, S. Jewson & C. Ziehmann. *Weather Derivative Modeling and Valuation*, Risk Books 2002.

[5] C.D. Brody, J. Syroka & M. Zervos "Dynamical Pricing of Weather Derivatives" *Quantitative Finance*, 2, pp. 189-198, 2002.

[6] R. Caballero, S. Jewson & A. Brix "Long Memory in Surface Air Temperature: Detection, Modeling and Application to Weather Derivative Valuation" *Climate Research*, 21, pp127-140, 2002.

[7] M. Cao & J. Wei "Pricing The Weather" *Risk,* 13 (5), 2000.

[8] M. Davis "Pricing Weather Derivatives by Marginal Value" *Quantitative Finance*, 1, pp.305-308, 2001.

[9] I. Daubechies. *Ten Lectures on Wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics)*, 1992.

[10] F. Dornier & M. Querel "Caution to the Wind" *Energy Power Risk Management*, pp. 30-32, 2000.

[11] S. Jewson & A. Brix *Weather Derivative Valuation*, Cambridge University Press, 2005.

[12] S.G. Mallat *A Wavelet Tour of Signal Processing.* San Diego: Academic Press, 1999.

[13] M. Moreno "Riding the Temp" *Futures and Options World*, 11, 2000

[14] P. Wojtaszczyk. *A Mathematical Introduction to Wavelets.* Cambridge: Cambridge University Press, 1997.

[15] A. Zapranis & A. Alexandridis "Weather Analysis & Weather Derivative Pricing" in proc. HFAA, Thessaloniki, December 2006.

[16] A. Zapranis & A. Alexandridis "Weather Derivatives Pricing: Modeling The Seasonal Residual Variance of an Ornstein-Uhlenbeck Temperature Process With Neural Networks" in proc. EANN, Thessaloniki, 29-31 August, 2007 (accepted, to appear).

[17] A. Zapranis & A.-P. Refenes, *Principles of Neural Model Identification, Selection and Adequacy: With Applications to Financial Econometrics*, Springer-Verlag, 1999.