

Utility-Based Pricing of the Weather Derivatives

Hélène Hamisultane ^{†*}

September 10, 2007
(revised version)

ABSTRACT

Since the underlying of the weather derivatives is not a traded asset, these contracts cannot be evaluated by the traditional financial theory. Cao and Wei (2004) price them by using the consumption-based asset pricing model of Lucas (1978) and by assuming different values for the constant relative risk aversion coefficient. Instead of taking this coefficient as given, we suggest in this paper to estimate it by using the consumption data and the quotations of one of the most transacted weather contracts which is the New York weather futures on the Chicago Mercantile Exchange (CME). We apply the well-known generalized method of moments (GMM) introduced by Hansen (1982) to estimate it as well as the simulated method of moments (SMM) attributed to Lee and Ingram (1991) and Duffie and Singleton (1993). This last method is studied since it is presumed to give satisfactory results in the case of the weather derivatives for which the prices are simulated. We find that the estimated coefficient from the SMM approach must have improbably high values in order to have the calculated weather futures prices matching the observations.

Keywords : Weather derivatives, Consumption-based asset pricing model, generalized method of moments, simulated method of moments, Monte-Carlo simulations, periodic variance.

JEL classification : C15, C22, C51, G13

[†] EconomiX , 200 avenue de la République, 92000 Nanterre, France. E-mail : helene.hamisultane@free.fr

* I would like to thank Sandrine Lardic and an economist of Sinopia for their useful comments.

1. Introduction

Weather impacts many sectors of the economy such as agriculture, construction, tourism and energy. It is estimated that 70% of the US economy is vulnerable to the unexpected weather conditions. To manage this risk, two power companies, Enron and Koch, negotiated the first weather derivative swap contract in 1997. The market has rapidly expanded to give birth to a standardized market, the Chicago Mercantile Exchange (CME), in 1999. Weather derivatives are financial instruments based on a weather index. They give a payment to the buyer of the contract when the underlying goes beyond a predetermined limit (the strike level). In contrast to the other derivatives, these contracts cannot be priced by using the traditional financial theory which consists in the elaboration of a hedging portfolio that attains the value of the weather derivative at the expiration date and by the non-arbitrage principle makes the price of the contract at time 0 equal to the cost of the hedging portfolio at time 0. Mathematically, this price corresponds to the discounted expectation of the payoff of the contingent claim under the equivalent martingale measure and not under the real probability measure. This pricing method is not possible here for the weather derivatives because the weather index is not a traded asset. It cannot be coupled with the riskless asset to form the hedging portfolio. Platen and West (2004) argue that the weather derivatives should be priced by the actuarial method since the weather index is not correlated with the growth optimal portfolio (GOP) which is approximated by the MCSI World index. Actuarial price is defined as the discounted expectation of the payoff with respect to the real probability plus a safety loading (or risk premium) payable by the insured. As noted by Platen and West (2004), this risk premium tends to diminish with the competition between insurers and when the reinsurance is available. Some authors such as Augros and Moréno (2002) and Brix, Jewson and Ziehmann (2002) emphasize the use of this pricing method for the weather derivatives because at the present time no hedging portfolio can be constituted and because the implementation of this method is simple. Indeed, the expectation of the payoff is calculated by averaging the number of simulations of the payoff and the price is then obtained by discounting at the riskless rate this expectation. But we can notice that this method does not account for the market information when quotations exist for the weather futures on the CME. Even if the organized market started in 1999, it is still nowadays illiquid but it is seen that some weather contracts have begun to be frequently traded (Jewson (2004)). Hamisultane (2006a) shows that incorporating the information contained in the prices of the most negotiated contracts can improve the predictions of the weather derivative values. Furthermore, Cao and Wei (2004) and Richards, Manfredo and Sanders (2004) find that the temperature is significantly correlated with the aggregate consumption and that the market price of weather risk is significantly different from zero. Therefore, this latter cannot be ignored for the valuation of these contracts and applying the actuarial approach in the case of the weather derivatives is not appropriate. Cao and Wei (2004) were the first to suggest the consumption-based asset pricing model of Lucas (1978) to price the weather derivatives. In this model, individuals seek to maximize a time-additive intertemporal discounted utility function that depends on a stochastic consumption. To calculate the prices, Cao and Wei (2004) and Richards *et al.* (2004) use different values postulated by the theory for the relative risk aversion coefficient but they do not estimate it from the data. In this paper, we propose to find it by using the weather futures prices since they are the only available weather derivative quotations on the market. We will concentrate on the New York contract which is one of the most transacted weather derivatives on the CME. Two estimation techniques will be conducted, the generalized method of moments (GMM) developed by Hansen (1982) and the simulated method of moments studied by Lee and Ingram (1991) and Duffie and Singleton (1993). The second method can be regarded as a calibration rather than an estimation method.

It is more complex to implement than the GMM because it necessitates for the simulations the modelling and the estimation of the processes that enter into the stochastic equilibrium model whereas the GMM only needs to use the available data of the variables in the model to determine the value of the risk aversion parameter. The SMM approach is considered here because the price of the weather derivatives cannot have a closed-form expression and therefore needs to be calculated by using the Monte-Carlo simulations. Then, we think that a risk aversion coefficient which is valued by calibration such as the simulated prices are the nearest to the observed prices should allow us to have more adequately reproduced prices. Our findings confirm that. For deriving this coefficient by the SMM procedure, we will model on one hand, the temperature and on the other hand, the consumption behavior. Weather derivatives can be based on rain, snow or frost index. We choose to centre our analysis on the temperature-based contracts which represent the majority of the negotiated contracts on the market. We will review all the models that have been suggested for capturing the features of the temperature behavior to select the proper one for the New York daily average temperature. We find that the risk aversion coefficient implied by the SMM way is very high and confirms what was put in evidence by the previous studies, that is, the incapacity of the consumption-based asset pricing model to generate plausible values. Mehra and Prescott (1985) showed that the model predicted an equity premium (the excess average return on stocks over the return on short-term T-bills) that was too small compared to the US annual 6% of the observed one while Weil (1989) pointed out that the predicted risk-free rate was too high. These two stylised facts are known as the “equity premium puzzle” and the “risk-free rate puzzle”.

The contribution of this paper is threefold : firstly, it estimates the risk aversion coefficient instead of taking it as given, secondly, it suggests the use of the SMM approach to estimate this coefficient and thirdly, it exposes the econometric proceeding that leads to the selection of the proper process for the daily average temperature. The structure of the paper is as follows: section 2 outlines the consumption-based asset pricing model for the weather derivatives, section 3 describes the estimation of the relative risk aversion coefficient by the GMM approach, section 4 comments the estimation of the coefficient by the SMM procedure, section 5 investigates the modelling of the daily average temperature after reviewing the literature on this topic, section 6 discusses the modelling and the estimation of the consumption growth rate process, section 7 involves the computation of the New York weather derivative prices by using on one hand the risk aversion parameter given by the GMM and on the other hand, the parameter induced by the SMM, the test of the overidentifying restrictions (or goodness-of-fit test) will also be performed and section 8 concludes the paper.

2. Theoretical framework

The model of Lucas (1978) considers an exchange economy where there is no money effect and a representative agent who maximizes the expected value of his or her time-additive utility function given by

$$E \left[\sum_{t=0}^{\infty} \beta^t U(C_t) \mid F_0 \right], \quad 0 < \beta < 1 \quad (1)$$

under the following budget constraint (we use here the Hansen and Singleton (1982)’s notation) :

$$C_t + P_t Q_t \leq K_t Q_{t-M} + W_t \quad (2)$$

where C_t is the real per-capita consumption at time t , $E[.]$ represents the expectation operator, F_0 corresponds to the information available to the agent at time 0, β is a discount factor, $U(.)$ is a strictly concave utility function with $U'(\cdot) > 0$ and $U''(\cdot) < 0$, P_t is the ex-dividend real price of the asset, Q_t stands for the quantity of the asset, K_t refers to post-dividend real price given by $K_t = (P_t + D_t)$, D_t is the real dividend and W_t denotes the real labor income at time t .

The first-order condition (called the ‘‘Euler condition’’) of this maximization problem given by (1) and (2) is defined as

$$E \left[\beta^M \frac{U'(C_{t+M})}{U'(C_t)} \frac{K_{t+M}}{P_t} \mid F_t \right] = 1. \quad (3)$$

In most of the studies, a constant relative risk aversion utility function is employed (see Mehra (2003) for explanations of this choice). It has the following form:

$$U(C_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma} \quad (4)$$

where $\gamma \geq 0$ is the relative risk aversion parameter which embeds the risk-neutral case when $\gamma = 0$. Putting this function in Eq.(3), we obtain

$$E \left[\beta^M \left(\frac{C_{t+M}}{C_t} \right)^{-\gamma} \frac{K_{t+M}}{P_t} \mid F_t \right] = 1. \quad (5)$$

By using Eq.(5), Cao and Wei (2004) show that the weather call option price $f(t, t_m)$ and the weather futures price $F(t, t_m)$ at time t for the maturity at date t_m are determined as follows :

$$f(t, t_m) = E \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} \max(I-K, 0) \mid F_t \right] \quad (6)$$

and

$$F(t, t_m) = E \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} I \mid F_t \right] \quad (7)$$

where K is the strike level and I denotes the temperature index at the maturity date. At the date t_m , we have $f(t_m, t_m) = \max(I-K, 0)$ and $F(t_m, t_m) = I$.

This temperature index can be the accumulation of the degree-days or the daily average temperatures. The average temperature for day t is defined as

$$T_t = \frac{T_t^{\max} + T_t^{\min}}{2} \quad (8)$$

where T_t^{\max} and T_t^{\min} represent respectively the maximum and minimum temperature at time t . The heating and cooling degree-day (HDD and CDD) for day t are expressed as follows

$$HDD_t = \max(65^\circ\text{F} - T_t, 0) \quad (9)$$

for October, November, December, January, February, March and April
and

$$CDD_t = \max(T_t - 65^\circ\text{F}, 0) \quad (10)$$

for May, June, July, August and September

where $65^\circ\text{F} \approx 18^\circ\text{C}$ is the reference temperature above which people start turning their air conditioners on for cooling and under which people start heating their homes.

The HDD and CDD indexes over a period of m days are then formulated as

$$I_m^H = \sum_{t=1}^m HDD_t \quad (11)$$

and

$$I_m^C = \sum_{t=1}^m CDD_t. \quad (12)$$

Cao and Wei (2004) calculate the weather derivative prices by taking three cases for the risk aversion coefficient : $\gamma = -2$, $\gamma = -10$ and $\gamma = -40$ ⁽¹⁾. They do not estimate it from the data. To fill the gap, we suggest in this paper to find the value of the risk aversion coefficient by using the consumption data and the weather futures prices since there is no available price for the weather options. The estimated parameter is thereafter employed to compute the weather futures prices. These daily prices are calculated by using the monthly consumption data which are disaggregated by a linear interpolation in order to obtain daily values. To derive the estimate of the risk aversion coefficient in a nonlinear expectation model, the GMM approach proposed by Hansen (1982) is the commonly applied procedure. However, the weather derivative price cannot be expressed in an analytical way and is generally calculated by the Monte-Carlo simulations. This point lets us expect that the estimated risk aversion coefficient from the SMM proceeding can give better reproduced prices than the GMM. The fact that Eq.(6) and Eq.(7) are formulated in terms of accumulated HDDs or CDDs does not allow us to find a closed-form representation for the price. Indeed, in the case of the HDD index, the expressions $E\left[\max\left(\sum_{t=1}^m \max(65^\circ\text{F}-T_t, 0) - K, 0\right) \mid F_t\right]$ and $E\left[\sum_{t=1}^m \max(65^\circ\text{F}-T_t, 0) \mid F_t\right]$ cannot be simplified. We show in section 7 that the estimates from the calibration method yield indeed good results but the obtained values of the risk aversion coefficient are very high and implausible from a theoretical point of view as numerous previous works relative to the consumption-based asset pricing have already mentioned (Mehra and Prescott (1985), Mankiw and Zeldes (1991), Ferson and Constantinides (1991) and Constantinides and Duffie (1996)). Before commenting these results, we first introduce and compare the two mentioned estimation methods of the risk aversion parameter.

3. Estimation of the constant relative risk aversion coefficient : the GMM approach

In order to estimate the constant relative risk aversion coefficient by the GMM procedure and with the weather futures prices, Eq.(7) must be turned into an unconditional expectation. By applying the law of iterated expectations, we then obtain :

$$E \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} \frac{F(t_m, t_m)}{F(t, t_m)} - 1 \right] = 0 \quad (13)$$

and

$$E[h_t(\theta)] = 0 \quad (14)$$

where $\theta = (\beta, \gamma)$ and $h_t(\theta) = \beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} \frac{F(t_m, t_m)}{F(t, t_m)} - 1$ for $t=1, \dots, n$. The term $h(\theta)$ can be considered as a disturbance.

The GMM estimators of β and γ consist in minimizing this quadratic form

$$Q_n(\theta) = \bar{h}(\theta)' W_n \bar{h}(\theta) \quad (15)$$

where $\bar{h}(\theta) = \frac{1}{n} \sum_{t=1}^n h_t(\theta)$, n is the number of observations of $h_t(\theta)$ and W_n is a positive definite symmetric nonsingular weighting matrix.

Since we only have one moment condition here (because we have assumed only one asset in the economy) and two unknowns, we cannot estimate these coefficients. To have at least as many moment conditions as unknowns, instrumental variables are used. Hansen and Singleton (1982) suggest to employ the lagged values of the consumption growth rate and return.

We rewrite Eq.(14) by introducing these instrumental variables to obtain the following orthogonality conditions :

$$E[h_t(\theta) \otimes z_t] = 0_L \quad (16)$$

And

$$E[g_t(\theta)] = 0_L \quad (17)$$

where $g_t(\theta) = h_t(\theta) \otimes z_t$, \otimes represents the Kronecker product, z_t refers to the vector of $Z \times 1$ instruments, 0_L is the $L \times 1$ vector of 0s, $L = 1 \times Z$ (we multiply Z to 1 because we have one variable $h_t(\theta)$ corresponding to the only one asset) and $z_t = \left(1, \left(\frac{C_{t_m}}{C_t} \right)_{-1}, \left(\frac{C_{t_m}}{C_t} \right)_{-2}, \dots, \left(\frac{F(t_m, t_m)}{F(t, t_m)} \right)_{-1}, \left(\frac{F(t_m, t_m)}{F(t, t_m)} \right)_{-2}, \dots \right)$.

Therefore, the quadratic form to minimize turns to :

$$Q_n(\theta) = \bar{g}(\theta)' W_n \bar{g}(\theta) \quad (18)$$

where $\bar{g}(\theta) = \frac{1}{n} \sum_{t=1}^n g_t(\theta)$, $\bar{g}(\theta)$ is a $L \times 1$ vector and W_n corresponds to a $L \times L$ matrix.

Andersen and Sørensen (1996) show that the just-identified case when there are as many moment conditions as unknowns produces poor results in terms of the root mean square error (RMSE) compared to those yielded by the overidentified case. For this reason, we will only consider the last configuration here.

When the system is overidentified, we cannot find solutions so that the sample moments of the observed data is equal to zero, i.e., $\sum_{t=1}^n g_t(\hat{\theta}) = 0$ where $\hat{\theta}$ corresponds to the estimated value of θ . We can only find solutions giving sample moments close to zero. A test allows us to check whether the moments are correctly specified. The test statistic is given by

$$J = n \times \hat{Q}_n(\hat{\theta}). \quad (19)$$

The J-statistic follows a χ^2_{L-p} distribution when the moment conditions are correctly specified where L corresponds to the number of conditions and p denotes the number of unknowns. If J is greater than the critical value from the χ^2_{L-p} distribution, the null hypothesis that the moments are not misspecified is rejected.

To minimize the quadratic form, we have to choose a weighting matrix. The choice of this matrix is important when the number of moment conditions is greater than the number of unknowns. As noted by Andersen and Sørensen (1996), an inappropriate choice may induce hard convergence to the solution. Hansen (1982) suggests to take $W_n = S^{-1}$ where S is the variance-covariance matrix of the random variable $g_t(\theta)$ so as to minimize the asymptotic covariance matrix Σ of the estimator $\hat{\theta}$. Under the assumption that $\{g_t(\theta)\}$ is a stationary⁽²⁾ and ergodic process, $\hat{\theta}$ is consistent and asymptotically normal (Hansen (1982)) : $n^{1/2}(\hat{\theta} - \theta) \rightsquigarrow N(0, \Sigma)$ and the matrix Σ is given by

$$\Sigma = (D'S^{-1}D)^{-1} \quad (20)$$

where $D = E\left(\frac{\partial g_t(\hat{\theta})}{\partial \hat{\theta}'}\right)$ is a matrix of dimension $L \times p$. D is estimated by

$$\hat{D} = \frac{1}{n} \sum_{t=1}^n \frac{\partial g_t(\hat{\theta})}{\partial \hat{\theta}'}. \quad (21)$$

The matrix S is defined as

$$S = \frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n E[g_t(\theta)g_s(\theta)']. \quad (22)$$

By change of indexes, we obtain

$$S = \frac{1}{n} \sum_{j=-n+1}^{n-1} \sum_{t=j+1}^n E[g_t(\theta)g_{t-j}(\theta)'] \text{ for } j \geq 0. \quad (23)$$

As $n \rightarrow \infty$, S is equal to 2π times the spectral density of $g_t(\theta)$ at frequency zero ($\lambda \approx 0$) given by

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} e^{-ij\lambda} E(g_t(\theta) g_{t-j}(\theta)'). \quad (24)$$

Then, the matrix S is usually approximated by this spectral density. Its estimator is expressed as⁽³⁾

$$\hat{S} = \Omega_0 + \sum_{j=1}^{n-1} [\Omega_j + \Omega_j'] \quad (25)$$

where

$$\Omega_j = \frac{1}{n} \sum_{t=j+1}^n g_t(\hat{\theta}) g_{t-j}(\hat{\theta})'. \quad (26)$$

As noted by Newey and West (1987), this estimator may not be positive definite in any finite sample when the number of sample autocovariances is not zero. They suggest the following consistent, positive semi-definite, heteroskedasticity and autocorrelation covariance (HAC) matrix estimate :

$$\tilde{S} = \Omega_0 + \sum_{j=1}^b \text{ker}(j) [\Omega_j + \Omega_j'] \quad (27)$$

where b stands for the bandwidth, $\text{ker}(j) = 1 - \frac{j}{b+1}$ is a weight and it is often called the Bartlett kernel function. Other kernel functions can also be considered such as the Parzen and the quadratic spectral kernel. We limit here our study to the Bartlett case since this latter performs well in terms of RMSE (see Andersen and Sørensen (1996) and Chiang and Kao (2005)).

The value of the bandwidth b can be fixed independently of the size of the sample or according to it. When it is data-dependent, it is determined by the Newey and West (1994)'s approach or by the Andrews (1991)'s method.

For the Newey and West (1994)'s approach, b is calculated as :

$$b = 1.1447 (\hat{\gamma} n)^{1/3} \quad (28)$$

where

$$\hat{\gamma} = \left(\frac{s_1}{s_0} \right)^2 \quad (29)$$

and

$$s_1 = 2 \sum_{j=1}^d j \hat{\sigma}_j, \quad s_0 = \hat{\sigma}_0 + 2 \sum_{j=1}^d \hat{\sigma}_j \quad (30)$$

with

$$d = 4 \left(\frac{n}{100} \right)^{2/9}, \quad \hat{\sigma}_j = \frac{1}{n-1} \sum_{t=j+2}^n g_t(\hat{\theta}) g_{t-j}(\hat{\theta})' \quad \text{and } j = 0, \dots, d \quad (31)$$

For the Andrews (1991)'s method, b is defined as :

$$b = 1.1447 (\hat{\alpha}_n)^{1/3} \quad (32)$$

where

$$\hat{\alpha} = \sum_{\ell=1}^L \frac{4 \hat{\rho}_\ell^2 \hat{\rho}_\ell^4}{K(1-\hat{\rho}_\ell)^6(1+\hat{\rho}_\ell)^2} \quad (33)$$

and

$$K = \sum_{\ell=1}^L \frac{\hat{\sigma}_\ell^4}{(1-\hat{\rho}_\ell)^4} \quad (34)$$

with $(\hat{\rho}_\ell, \hat{\sigma}_\ell)$ denoting the estimates of the autoregressive and innovation variance parameter for each of the moments $(\ell = 1, \dots, L)$.

For our study, we will employ both the Newey and West's (1994) and Andrews (1991)'s methods since we have no intuition about what value to choose for the bandwidth.

Andrews and Monahan (1992) show that the prewhitening procedure can improve the performance of the kernel HAC estimator. It consists in fitting a b^{th} order vector autoregressive (VAR) to $g_t(\hat{\theta})$. We write here the VAR(1) form for simplicity :

$$g_t(\hat{\theta}) = \hat{B} g_{t-1}(\hat{\theta}) + g_t^*(\hat{\theta}). \quad (35)$$

Thereafter, the prewhitened residuals $g_t^*(\hat{\theta})$ are used to compute the kernel HAC estimator \tilde{S}^* which is next "recolored" to obtain the VAR prewhitened kernel estimator \tilde{S}_{pw} :

$$\tilde{S}_{pw} = (I - \hat{B})^{-1} \tilde{S}^* (I - \hat{B}')^{-1}. \quad (36)$$

The matrix \tilde{S}_{pw} is then employed to minimize the quadratic function (18).

We would like to point out that when starting the implementation of the GMM techniques, consistent values of $\hat{\theta}$ are required to compute the estimator of S but these latter are unknown. They can be obtained by using $W_n = I_n^{-1}$ in the first step where I_n refers to the identity matrix since, according to Hansen (1982), any weighing matrix can give consistent estimators. In the second step, the estimates of θ are used to form $W_n = S^{-1}$.

4. Calibration of the constant relative risk aversion coefficient : the SMM approach

In the GMM approach, the parameters β and γ are estimated from the data which are observed and not simulated. Therefore, these estimates can give no satisfactory results in the case where Monte-Carlo simulations are used to compute the prices. To find the values of β and γ such as the simulated prices are the closest to the observations, we use the SMM approach which consists in solving the following optimization problem in the weather derivative case :

$$\underset{\beta, \gamma}{\text{Min}} Q_N(\beta, \gamma) = \left(E \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} I \mid F_t \right] - F(t, t_m) \right)' W_N \left(E \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} I \mid F_t \right] - F(t, t_m) \right) \quad (37)$$

\Leftrightarrow

$$\underset{\beta, \gamma}{\text{Min}} Q_N(\beta, \gamma) = \left(\frac{1}{N} \sum_{t=1}^N \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} I \right] - F(t, t_m) \right)' W_N \left(\frac{1}{N} \sum_{t=1}^N \left[\beta^{(t_m-t)} \left(\frac{C_{t_m}}{C_t} \right)^{-\gamma} I \right] - F(t, t_m) \right) \quad (38)$$

where N represents the number of simulations and $I = I_m^H = \sum_{t=1}^m \max(65^\circ F - T_t, 0)$ or $I = I_m^C = \sum_{t=1}^m \max(T_t - 65^\circ F, 0)$.

On the contrary to the earlier method, no instrumental variable is needed here because we have more moment conditions than unknowns. Indeed, we have as many moment conditions as observed quotations for $F(t, t_m)$ to estimate β and γ . In that case, we can conduct the overidentifying restrictions test which was exposed before to check the goodness of fit of the model. Unlike the GMM principle, the present method requires the specification of the consumption and temperature processes in order to simulate them. In the following sections, we will pay attention to the modelling of the movements of the temperature and consumption. Aside from these differences, the implementation of the SMM techniques is based on the same procedures as those presented in the GMM approach with n replaced by N .

5. Modelling the temperature behavior

To find a model that captures the characteristics of our temperature series, we first review the main processes that were suggested in literature. In light of observations, Cao and Wei (2004) list the following features of the daily average temperature behavior. It reverts to a mean which has a seasonal cyclical pattern, it tends to rise with time due to a global warming trend, it has a larger variation in the winter than in the summer and the driving noise is Gaussian. They suggest the following process

$$T_t = m_t + s_t + \sum_{j=1}^p \rho_j U_{t-j} + \sigma_t \tilde{\varepsilon}_t, \quad \tilde{\varepsilon}_t \sim \text{iid}(0,1) \quad (39)$$

$$m_t = \frac{\beta_0}{365} \left(t - \frac{n}{2} \right), \quad (40)$$

$$s_t = \bar{T}_t^y = \bar{T}_d (365 \times 1) \otimes \ell_{(Y \times 1)}, \quad (41)$$

$$\bar{T}_d = \frac{1}{Y} \sum_{y=1}^Y T_{d,y}, \quad (42)$$

where ℓ is a $Y \times 1$ vector of ones, $d=1, \dots, 365$ corresponds to the day of the year $y=1, \dots, Y$, $t=1, \dots, n$ where $n=365 \times Y$ (February 29 is removed for leap-years),

$$U_t = T_t - \left(\frac{\beta_0}{365} \left(t - \frac{n}{2} \right) + \bar{T}_t^y \right) \quad (43)$$

and

$$\sigma_t = \sigma_0 - \sigma_1 \left| \sin \left(\frac{\pi t}{365} + \varphi \right) \right|. \quad (44)$$

In order to analyse the significativity of the trend alone in Eq.(40), we rewrite this equation as follows :

$$m_t = \frac{\beta_1}{365} t - \frac{\beta_0}{365} \frac{n}{2}. \quad (45)$$

Roustant (2002) proposes a model which incorporates a smoother cyclical component for the mean and for the variance. Furthermore, he allows the driving noise to have an ARMA(p,q) (AutoRegressive Moving Average) structure instead of an AR process like in Cao and Wei (2004), p and q representing respectively the orders of the AR and MA parts. His model is expressed as :

$$T_t = m_t + s_t + \sum_{j=1}^p \rho_j U_{t-j} - \sum_{l=1}^q \Theta_l \varepsilon_{t-l} + \varepsilon_t \quad (46)$$

which can be rewritten in that way

$$[1 - \rho(L)] (T_t - m_t - s_t) = [1 - \Theta(L)] \varepsilon_t, \quad (47)$$

where

$$\rho(L) = \sum_{j=1}^p \rho_j L^j, \quad \Theta(L) = \sum_{l=1}^q \Theta_l L^l \quad (48)$$

with $L^n y_t = y_{t-n}$,

$$U_t = T_t - m_t - s_t, \quad (49)$$

$$m_t = d t + e , \quad (50)$$

$$s_t = a_1 \cos(\omega t) + b_1 \sin(\omega t) + a_2 \cos(2\omega t) + b_2 \sin(2\omega t) , \quad (51)$$

$$\omega = \frac{2\pi}{365} , \quad (52)$$

$$\varepsilon_t = \sigma_t \tilde{\varepsilon}_t , \quad \tilde{\varepsilon}_t \sim \text{iid}(0,1) \quad (53)$$

and

$$\sigma_t = a + b \cos(\omega t) + c \sin(\omega t). \quad (54)$$

Campbell and Diebold (2004) introduce a more generalized model including a GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) process in the variance. It is formulated as follows :

$$T_t = m_t + s_t + \sum_{j=1}^p \rho_j T_{t-j} + \sigma_t \tilde{\varepsilon}_t , \quad (55)$$

$$m_t = \sum_{m=0}^M \beta_m t^m , \quad (56)$$

$$s_t = \sum_{k=1}^K \left[\delta_{c,k} \cos\left(2\pi k \frac{d(t)}{365}\right) + \delta_{s,k} \sin\left(2\pi k \frac{d(t)}{365}\right) \right] \quad (57)$$

and ⁽⁴⁾

$$\sigma_t^2 = c + \sum_{q=1}^Q \left[\gamma_{c,q} \cos\left(2\pi q \frac{d(t)}{365}\right) + \gamma_{s,q} \sin\left(2\pi q \frac{d(t)}{365}\right) \right] + \sum_{r=1}^R \mu_r (\sigma_{t-r} \tilde{\varepsilon}_{t-r})^2 + \sum_{s=1}^S v_s \sigma_{t-s}^2 \quad (58)$$

where $d(t)$ is a repeating step function that cycles through $1, \dots, 365$. On the contrary to the simple GARCH process, positiveness restrictions on the coefficients μ and v are not necessary

here because the estimated part $c + \sum_{q=1}^Q \left[\gamma_{c,q} \cos\left(2\pi q \frac{d(t)}{365}\right) + \gamma_{s,q} \sin\left(2\pi q \frac{d(t)}{365}\right) \right]$ guaranties

the positiveness of the volatility.

Note that T_t is explained by m_t , s_t and lagged values of T_t . We think that this process should be written with lagged values of $(T_t - m_t - s_t)$ instead of T_t in order to obtain an ARMA representation (precisely an AR process here) which is the common way to model a time

series. Therefore, we suggest to estimate the following process that we call the "Campbell and Diebold's modified model" :

$$T_t = m_t + s_t + \sum_{j=1}^p \rho_k (T_{t-j} - m_{t-j} - s_{t-j}) + \sigma_t \tilde{\varepsilon}_t . \quad (59)$$

Benth and Šaltytė-Benth (2005) use a discretized version of the Ornstein-Uhlenbeck process. It is written like an AR(1). In their model, they do not consider the trend because this latter is not significant. Their sample size regroups 4846 observations while the preceding authors work with a larger dataset (7300 for Cao and Wei (2004) and 7670 for Roustant (2002)). Their model is stated as follows :

$$T_t = m_t + s_t + \rho(T_{t-1} - m_{t-1} - s_{t-1}) + \sigma_t \tilde{\varepsilon}_t , \quad (60)$$

$$m_t = a_0 \quad (61)$$

and

$$s_t = a_1 \cos\left(\frac{2\pi}{365}(t-t_0)\right). \quad (62)$$

They account for the seasonality in the variance by using a non parametric estimator, i.e.

$$\hat{\sigma}_t^2 = \bar{\hat{\varepsilon}}_d^2 (365 \times 1) \otimes \ell_{(Y \times 1)} , \quad (63)$$

$$\bar{\hat{\varepsilon}}_d^2 = \frac{1}{Y} \sum_{y=1}^Y \hat{\varepsilon}_{d,y}^2 \quad (64)$$

and

$$\hat{\varepsilon}_{d,y} = T_{d,y} - \hat{m}_{d,y} - \hat{s}_{d,y} - \hat{\rho}(T_{d,y-1} - \hat{m}_{d,y-1} - \hat{s}_{d,y-1}) \quad (65)$$

where \hat{m}_t , \hat{s}_t and $\hat{\rho}$ denote the estimated values.

Many other authors have also appealed to the continuous time processes for reproducing the observed temperatures. Brody, Syroka and Zervos (2002), Benth (2003) and Hamisultane (2006b) deal with the mean-reverting fractional Brownian motion process. Richards, Manfredo and Sanders (2004) and Hamisultane (2006a) consider respectively a mean-reverting jump-ARCH process and a mean-reverting jump-EGARCH (Exponential GARCH) process. The drawbacks of these models are they produce very volatile prices when using Monte-Carlo simulations (see Hamisultane (2006a) and Hamisultane (2006b)) and their AR(1) structure does not fit the observations. Based on the information criteria, we show in the following that the proper model for the temperature must encompass a higher lagged structure. This finding reveals the inability of the continuous time processes to capture the characteristics of the temperature. This point was also underlined in Roustant (2002). For these reasons, we will not examine these processes in this paper. However, we will keep the Benth and Šaltytė-Benth (2005)'s process for our study to gauge the precision of its results compared to those given by the other models.

Another feature of the temperature that has been mentioned in the literature is its long memory. Caballero, Jewson and Brix (2002), Moréno (2003) and Hamisultane (2006b) reveal the presence of a persistent phenomenon in the mean and also in the variance of the temperature. Hamisultane (2006b) shows that the use of the long memory process ARFIMA-FIGARCH (Auto Regressive Fractionally Integrated Moving Average - Fractionally Integrated GARCH) gives prices that are not very different from those induced by the short memory process AR-GARCH. Therefore, for the sake of simplicity, we do not consider here this characteristic of the temperature.

We now use all these elements to build a model for our series. We possess historical data of the New York daily average temperature for a period from 01/01/1993 to 12/31/2005 which represents 4745 observations (after dropping February 29). We start our investigation by the modelling of the deterministic part (m_t and s_t) of the temperature. We estimate the Campbell and Diebold's modified model which encompasses the Roustant (2002)'s and Benth and Šaltytė-Benth (2005)'s representation for the trend and seasonal component. We use the maximum likelihood method to estimate all the models. Based on the information criteria which are the Akaike (AIC), the Schwartz (SC) and the Hannan-Quinn (HQ) measures⁽⁵⁾, we select the orders $M=0$, $K=3$ and $p=5$. The results of the estimation of the part including $M=0$ and $K=3$ are reported in Table 1 while the estimates of $p=5$ are presented in Table 2. Our model does not include a trend like in Benth and Šaltytė-Benth (2005)'s specification. Moreover, the order of the seasonal part is larger than the Roustant (2002)'s and Benth and Šaltytė-Benth (2005)'s suggestions. When adding moving average (MA) terms, i.e., when writing the model as follows :

$$T_t = m_t + s_t + \sum_{j=1}^p \rho_j (T_{t-j} - m_{t-j} - s_{t-j}) - \sum_{l=1}^q \Theta_l (\sigma_{t-l} \tilde{\varepsilon}_{t-l}) + \sigma_t \tilde{\varepsilon}_t, \quad (66)$$

we do not decrease the information criteria as shown in Table 2. Our model has then a pure AR process for the disturbance. Next, we compare the results in Table 1 to the estimates of the trend and the seasonality of the the Cao and Wei 's model which are shown in Table 3. Since the trend was not significant, it does not appear in Table 3. After removing the seasonal part from the initial observations of the temperature to obtain a stationary time series, we model the behavior of the new sample by the AR process. The smallest values of the information criteria are reached for $p=5$. The results of the estimation are put in Table 4. We see that the Cao and Wei 's representation has information measures which are smaller than those of the Campbell and Diebold 's estimated model. When MA terms are introduced, they do not improve the results as we can notice in Table 3. We therefore choose the AR(5) process to represent the deterministic part of the New York temperature. We now check whether its residuals are white noises. Table 4 puts in evidence the non autocorrelation of the residuals since the corresponding H statistic of Durbin (in absolute value) is below 1.96 at 5% level. We use here the H statistic instead of the Durbin-Watson measure to avoid the bias due to the use of the autoregressive model. Table 5 sets out the ARCH test which shows that the variance of the residuals is not constant since the probability that the statistic nR^2 exceeds the value of 77.74 under the null hypothesis is close to zero (so inferior to the 5% level) which means that the alternative hypothesis that the variance is not homoskedastic is accepted. This finding involves the necessity of modelling the variance of the temperature. We estimate five possible representations for the variance which results are reported in Tables 6. Concerning the Campbell and Diebold's variance, we use the information measures to select the orders Q , R and S of the equation. We found that $Q= 1$, $R=1$ and $S=1$ is the best combinaison. It

minimizes the AIC, SC and HQ information criteria. Our results in Table 6 reveal that accounting only for the seasonality of the volatility is not enough to obtain a good model since the Cao and Wei 's and Roustant's variance processes fail to have information measures which are inferior to those of the Campbell and Diebold's suggestion. In the same way, considering only the GARCH structure is not an appropriate choice. Finally, the Campbell and Diebold's process appears to be a good mixture of all these elements but it does not represent the "best" model for the volatility of the New York temperature when we compare it to the Benth and Šaltytė-Benth's estimated model. Indeed, the information criteria of this latter are lower than those of the Campbell and Diebold's representation. We plot it in Figure 2. We depict also the Campbell and Diebold's variance model in Figure 1 for comparison. We see that this last process is too smooth to capture the irregularities of the observed squared residuals. In conclusion, we choose the estimated model of Cao and Wei for the trend and the seasonality and the process of Benth and Šaltytė-Benth for the volatility. This representation will be used to simulate the temperature in the implementation of the SMM approach. Now, we need to model the behavior of the consumption. This point is treated in the following section.

Table 1 : Estimation of the coefficients of the trend and the seasonality of the Campbell and Diebold's model

	Estimation	t-statistic
β_0	55.95	548.60
$\delta_{c,1}$	-20.39	-141.36
$\delta_{c,2}$	-0.22	-1.55
$\delta_{c,3}$	-0.02	-0.14
$\delta_{s,1}$	-9.16	-63.51
$\delta_{s,2}$	0.17	1.20
$\delta_{s,3}$	-0.96	-6.63
LL	-15979.96	
AIC	6.7384	
SC	6.7480	
HQ	6.7418	

LL represents the value of the log-likelihood at the optimum ,
AIC, SC and HQ refer respectively to the Akaike, Schwarz and Hannan-Quinn information measures,

$$T_t = m_t + s_t , \quad m_t = \sum_{m=0}^M \beta_m t^m \quad \text{and} \quad s_t = \sum_{k=1}^K \left[\delta_{c,k} \cos\left(2\pi k \frac{d(t)}{365}\right) + \delta_{s,k} \sin\left(2\pi k \frac{d(t)}{365}\right) \right].$$

Table 2 : Estimation of the coefficients of the ARMA process for the stationnarized series by the removal of the estimated trend and seasonality of the Campbell and Diebold's model

	AR process		ARMA process	
	Estimation	t-statistic	Estimation	t-statistic
ρ_1	0.80	55.31	1.37	7.00
ρ_2	-0.27	-14.32	-0.72	-4.57
ρ_3	0.13	7.09	0.28	5.05
ρ_4	-0.03	-1.59	-0.10	-3.21
ρ_5	0.04	2.87	0.04	2.40
Θ_1			-0.57	-2.90
LL	-14405.49		-14403.66	
AIC	6.0804		6.0800	
SC	6.0872		6.0882	
HQ	6.0828		6.0829	
H	$ -0.897 = 0.897$		NA ⁽⁶⁾	

LL represents the value of the log-likelihood at the optimum , AIC, SC and HQ refer respectively to the Akaike, Schwarz and Hannan-Quinn information measures, H stands for the H statistic of Durbin and

$$(T_t - m_t - s_t) = \sum_{j=1}^p \rho_k (T_{t-j} - m_{t-j} - s_{t-j}) - \sum_{l=1}^q \Theta_l \varepsilon_{t-l} + \varepsilon_t.$$

Table 3 : Estimation of the coefficients of the trend and the seasonality of the Cao and Wei's model

	Estimation	t-statistic
β_0	55.95	566.31
κ	1.00	161.09
LL	-15831.71	
AIC	6.6738	
SC	6.6766	
HQ	6.6748	

LL represents the value of the log-likelihood at the optimum , AIC, SC and HQ refer respectively to the Akaike, Schwarz and Hannan-Quinn information measures, $T_t = m_t + s_t$, $m_t = \frac{\beta_1}{365} t - \frac{\beta_0}{365} \frac{n}{2}$

and $s_t = \kappa (\bar{T}_d (365 \times 1) \otimes \ell_{(Y \times 1)})$ with $\bar{T}_d = \frac{1}{Y} \sum_{y=1}^Y T_{d,y}$ where n corresponds to the number of observations,

ℓ is a $Y \times 1$ vector of ones, $d=1, \dots, 365$ denotes the day of the year $y=1, \dots, Y$ and $t=1, \dots, n$.

Table 4 : Estimation of the coefficients of the ARMA process for the stationnarized series by the removal of the estimated trend and seasonnality of the Cao and Wei's model

	AR process		ARMA process	
	Estimation	t-statistic	Estimation	t-statistic
ρ_1	0.81	55.63	1.37	7.44
ρ_2	-0.26	-14.18	-0.72	-4.82
ρ_3	0.14	7.40	0.29	5.42
ρ_4	-0.03	-1.75	-0.10	-3.42
ρ_5	0.04	3.06	0.04	2.57
Θ_1			-0.57	-3.08
LL	-14227.20		-14225.40	
AIC	6.0051		6.0048	
SC	6.0120		6.0130	
HQ	6.0075		6.0077	
H	$ -1.500 = 1.500$		NA	

LL represents the value of the log-likelihood at the optimum , AIC, SC and HQ refer respectively to the Akaike, Schwarz and Hannan-Quinn information measures, H stands for the H statistic of Durbin and

$$(T_t - m_t - s_t) = \sum_{j=1}^p \rho_k (T_{t-j} - m_{t-j} - s_{t-j}) - \sum_{l=1}^q \Theta_l \varepsilon_{t-l} + \varepsilon_t.$$

Table 5 : ARCH test for the squared residuals (RESID²) of the Cao and Wei 's AR(5) process .

ARCH Test:

F-statistic	39.49447	Probability	0.000000
Obs*R-squared	77.74210	Probability	0.000000

Test Equation:
 Dependent Variable: RESID²
 Method: Least Squares
 Sample(adjusted): 8 4745
 Included observations: 4738 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	19.59371	0.699188	28.02353	0.0000
RESID ² (-1)	0.088121	0.014480	6.085564	0.0000
RESID ² (-2)	0.084879	0.014481	5.861505	0.0000

R-squared	0.016408	Mean dependent var	23.69354
Adjusted R-squared	0.015993	S.D. dependent var	36.45495
S.E. of regression	36.16226	Akaike info criterion	10.01454
Sum squared resid	6192004.	Schwarz criterion	10.01863
Log likelihood	-23721.45	F-statistic	39.49447
Durbin-Watson stat	2.006676	Prob(F-statistic)	0.000000

Table 6 : Estimation of the coefficients of the Cao and Wei 's AR(5) process and of the variance equation

	AR(5) – GARCH(1,1) (*)		AR(5) – variance of Cao and Wei (**)		AR(5) – variance of Roustant (***)
	Estimation		Estimation		Estimation
ρ_1	0.81 (54.40)	P_1	0.81 (57.72)	ρ_1	0.81 (56.43)
ρ_2	-0.26 (-14.27)	P_2	-0.26 (-14.32)	ρ_2	-0.26 (-14.41)
ρ_3	0.14 (7.12)	P_3	0.13 (6.98)	ρ_3	0.13 (7.02)
ρ_4	-0.03 (-1.70)	P_4	-0.03 (-1.77)	ρ_4	-0.03 (-1.84)
ρ_5	0.04 (2.94)	P_5	0.04 (2.76)	ρ_5	0.04 (2.84)
α_0	0.47 (3.52)	σ_0	6.57 (47.91)	a	4.80 (112.14)
α_1	0.05 (6.88)	σ_1	2.76 (15.42)	b	0.92 (13.91)
α_2	0.93 (91.96)	φ	2.80 (101.20)	c	0.74 (11.52)
LL	-14128.89	LL	-14104.62	LL	-14094.12
AIC	5.9649	AIC	5.9547	AIC	5.9503
SC	5.9758	SC	5.9656	SC	5.9612
HQ	5.9688	HQ	5.9585	HQ	5.9541
	AR(5) – variance of Campbell and Diebold (****)		AR(5) – variance of Benth and Šaltytė- Benth (****)		
	Estimation		Estimation		
ρ_1	0.81 (51.22)	ρ_1	0.82 (56.41)		
ρ_2	-0.27 (-13.40)	ρ_2	-0.26 (-11.37)		
ρ_3	0.13 (6.76)	ρ_3	0.15 (8.52)		
ρ_4	-0.03 (-1.62)	ρ_4	-0.05 (-2.10)		
ρ_5	0.04 (2.83)	ρ_5	0.05 (3.08)		
c	5.58 (36.04)	η	1.00 (55.18)		
$\gamma_{c,1}$	2.15 (2.95)				
$\gamma_{s,1}$	1.45 (2.08)				
μ_1	0.05 (4.08)				
v_1	0.72 (43.95)				
LL	-14078.52	LL	-13860.78		
AIC	5.9445	AIC	5.8510		
SC	5.9582	SC	5.8591		
HQ	5.9493	HQ	5.8538		

LL represents the value of the log-likelihood at the optimum , AIC, SC and HQ refer respectively to the Akaike, Schwarz and Hannan-Quinn information measures, the numbers in parentheses indicate the value of the t-statistic.

(*) $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$ where $\alpha_0 > 0$, $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$

(**) $\sigma_t = \sigma_0 - \sigma_1 \left| \sin\left(\frac{\pi t}{365} + \varphi\right) \right|$

(***) $\sigma_t = a + b \cos(\omega t) + c \sin(\omega t)$ where $\omega = \frac{2\pi}{365}$

(****) $\sigma_t^2 = c + \sum_{q=1}^Q \left[\gamma_{c,q} \cos\left(2\pi q \frac{d(t)}{365}\right) + \gamma_{s,q} \sin\left(2\pi q \frac{d(t)}{365}\right) \right] + \sum_{r=1}^R \mu_r \varepsilon_{t-r}^2 + \sum_{s=1}^S v_s \sigma_{t-s}^2$

(*****) $\sigma_t^2 = \eta (\hat{\varepsilon}_d^2)_{(365 \times 1)} \otimes \ell_{(Y \times 1)}$ where $\hat{\varepsilon}_d^2 = \frac{1}{Y} \sum_{y=1}^Y \hat{\varepsilon}_{d,y}^2$, $\hat{\varepsilon}_{d,y} = T_{d,y} - \hat{m}_{d,y} - \hat{s}_{d,y} - \hat{\rho}(T_{d,y-1} - \hat{m}_{d,y-1} - \hat{s}_{d,y-1})$,

\hat{m}_t, \hat{s}_t and $\hat{\rho}$ denote the estimated values , ℓ is a $Y \times 1$ vector of ones, $d=1, \dots, 365$ corresponds to the day of the year $y=1, \dots, Y$, $t=1, \dots, n$ and $n = Y \times 365$ represents the number of observations.

Figure 1 : Squared residuals of the Cao and Wei's AR(5) process and estimated variance of Campbell and Diebold.

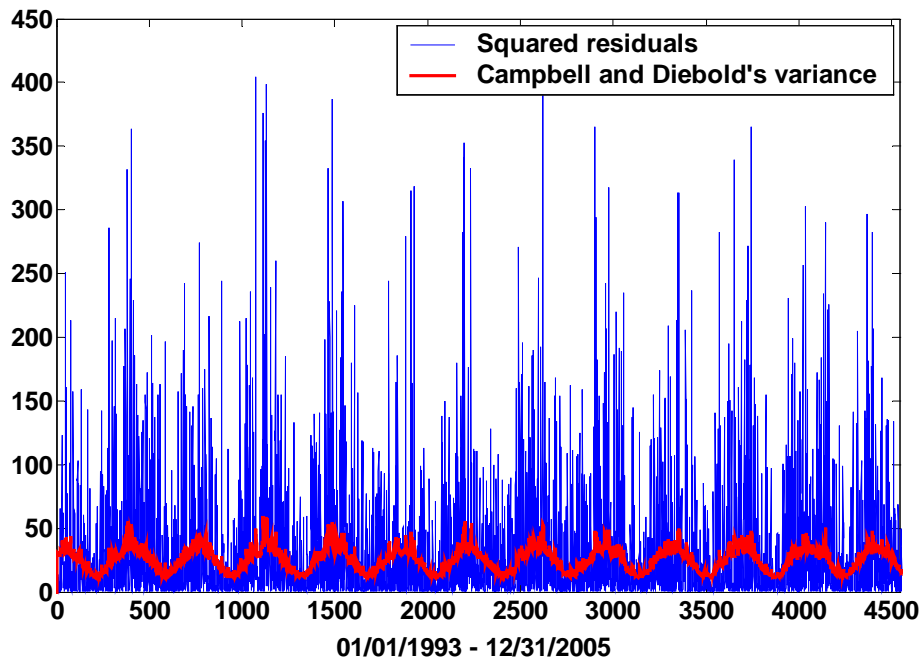
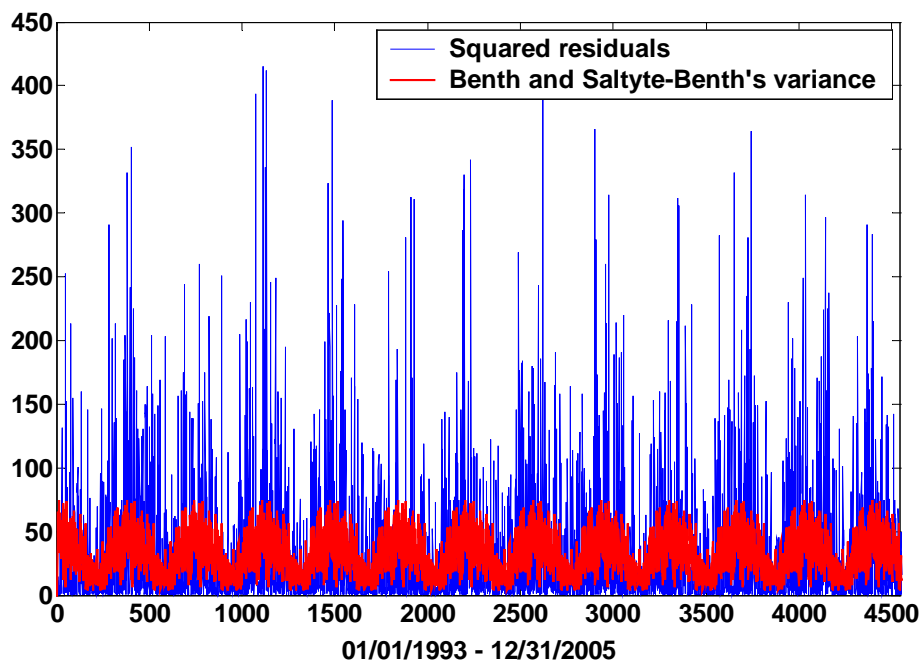


Figure 2 : Squared residuals of the Cao and Wei's AR(5) process and estimated variance of Benth and Šaltytė-Benth .



6. Modelling the behavior of the growth rate of consumption

We notice that the Euler condition is expressed in terms of the growth rate of the real per-capita consumption $\frac{C_{t_m}}{C_t}$. To make the simulation and estimation of the process of this variable more practical, the Euler condition is reformulated as follows :

$$E \left[\beta^{(t_m-t)} e^{-\gamma(\ln C_{t_m} - \ln C_t)} I \mid F_t \right] = F(t, t_m) \quad (67)$$

Following Cao and Wei (2004), we assume that $\ln C_t$ follows a Markov process stated as :

$$\ln C_t = \alpha_0 + \alpha_1 \ln C_{t-1} + \mu_t \quad (68)$$

where $\alpha_1 \leq 1$, $\mu_t = \sqrt{1-\rho^2} \sigma_v \tilde{v}_t + \sigma_t \left[\rho \tilde{\varepsilon}_t + \eta_1 \tilde{\varepsilon}_{t-1} + \dots + \eta_m \tilde{\varepsilon}_{t-m} \right]$, σ_v is the volatility of the noise which is independent of the temperature disturbance, σ_t refers to the volatility of the temperature, ρ stands for the correlation coefficient between $\ln C_t$ and the temperature, η_1, \dots, η_m are the coefficients of the lagged innovations of the temperature which impact the consumption, $\tilde{v}_t \sim \text{iid}(0,1)$ and $\tilde{\varepsilon}_t \sim \text{iid}(0,1)$.

Cao and Wei (2004) use monthly total personal consumption data which were taken from the Federal Reserve Bank of Saint-Louis. However, Mankiw and Zeldes (1991) point out that the use of the total consumption is not correct since it includes the stockholder and nonstockholder consumption. This latter is unlikely to satisfy the Euler condition of the consumption-based asset pricing. Using the data provided by the Panel Study of Income Dynamics (PSID), Mankiw and Zeldes (1991) isolate the families who hold stocks and find a value of the risk aversion coefficient which is less important than in the case where the total consumption is accounted for since the consumption of this group is more correlated with the stock market. This finding does not resolve yet the equity premium puzzle raised by Mehra and Prescott (1985) since the value of the risk aversion is still too high but it constitutes an encouraging result. For our study, we cannot use these data because their frequency is a year while we want to calculate the weather derivative price for a day. The conversion of these data into a daily series by a linear interpolation will not give satisfactory results. We therefore choose to work with the monthly consumption data for which the disaggregation will not produce a too imprecise daily series. Unlike Cao and Wei (2004), we use expenditures on nondurable goods and services in order to have an utility function which is time-separable (Ferson and Constantinides (1991)). These data series are available in nominal and real measures from the Federal Reserve Bank of Saint-Louis. For our framework, we choose the real series. In order to obtain the per-capita consumption, we divide the expenditures on nondurables and services by the population. We need not here to deflate the weather futures prices that enter into the Euler condition since they are quoted in degree-day index points and not in value. For estimating the process of the daily real per-capita consumption, we use

observations for a period spanning from 01/01/1993 to 12/31/2005. We then obtain the following result :

$$\ln C_t = -0.000560 + 1.000062 \ln C_{t-1} - 2.14 \times 10^{-7} \text{res}_t + \sqrt{1 - (-2.14 \times 10^{-7})^2} \times \sigma_v \tilde{v}_t \quad (69)$$

(-3.40) (60073.28) (-0.81)

where the number in parentheses refers to the t-statistic and res_t is the residual at time t of the Cao and Wei 's AR(5) process, it represents the estimation of the innovation $\sigma_t \tilde{\xi}_t$ of the temperature.

The standard deviation σ_v in equation (69) is estimated by :

$$\sigma_v = \sqrt{\frac{\hat{v}_t^2}{n-k}} \times \frac{1}{\sqrt{1 - (-2.14 \times 10^{-7})^2}} \quad (70)$$

where $\hat{v}_t = \ln C_t - (-0.000560 + 1.000062 \ln C_{t-1} - 2.14 \times 10^{-7} \text{res}_t)$, n is the number of observations and k is the number of dependent variables (we have here three variables).

Contrary to Cao and Wei (2004), our correlation coefficient is not significant (the t-statistic is below 1.96 at 5% level). We think that this is may be due to the choice of the linear interpolation for the consumption and decide to convert the daily temperature data into a monthly series as did Cao and Wei (2004) in order to compute the correlation coefficient between the monthly temperature and consumption. We see in Table 7 that it is still non significant even if we split the observations into HDD and CDD periods. We also use the total consumption expenditures like in Cao and Wei (2004). The results are displayed in Table 8. We choose the real and nominal measures for this aggregate and even in those cases, it remains non significant. We think that this difference comes from the fact that Cao and Wei (2004) use longer time series than ours (240 monthly observations versus 156 for us). Moreover, they average the temperatures across five cities whereas we only employ the New York temperature. We decide to keep the relation given in Eq.(69) for our simulations even if the correlation between the New York temperature and the consumption on nondurables and services is not proved since assuming no link between these two variables is equivalent to saying that the agents are risk neutral in the case of the weather derivatives ($\gamma=0$), that is to say that no market price of risk is required for these constructs which is inconsistent with our intuition (weather derivatives cannot be hedged by a self-financing portfolio, therefore a risk premium is expected) and with the findings of the previous works (Cao and Wei (2004) and Richards *et al.* (2004)).

After modelling the consumption and the temperature behavior, we can implement the SMM approach to find the value of the risk aversion coefficient. This procedure as well as the GMM methodology are run in the following section.

Table 7 : Estimation of the log-consumption process with US real per-capita consumption expenditures on nondurables and services

	Daily data			Monthly data		
	HDD period (n=2756)	CDD period (n=1989)	All days (n=4745)	HDD period (n=91)	CDD period (n=65)	All months (n=156)
α_0	-0.0007 (-0.49)	0.0008 (0.25)	-0.0006 (-3.40)	-0.0088 (-0.16)	0.0295 (0.28)	-0.0068 (-0.25)
α_1	1.0001 (6577.28)	0.9999 (3230.52)	1.0001 (60073.28)	1.0012 (177.33)	0.9974 (94.08)	1.0008 (362.24)
ρ	-9.17×10^{-7} (-0.42)	-6.15×10^{-6} (-1.04)	-2.35×10^{-7} (-0.88)	3.01×10^{-5} (0.07)	-0.0004 (-0.36)	-0.0002 (-0.74)

HDD period regroups January, February, March, April, October, November, December, CDD period spans from May to September, n denotes the number of observations and the figure in parentheses refers to the t-statistic.

Table 8 : Estimation of the log-consumption process with US total monthly consumption expenditures

	Real data			Nominal data		
	HDD period (n=91)	CDD period (n=65)	All months (n=156)	HDD period (n=91)	CDD period (n=65)	All months (n=156)
α_0	0.0535 (0.34)	0.1522 (0.56)	0.0266 (0.36)	0.0104 (0.07)	0.0555 (0.21)	0.0016 (0.03)
α_1	0.9983 (188.43)	0.9951 (108.60)	0.9992 (400.80)	0.9999 (209.15)	0.9985 (113.91)	1.0001 (573.33)
ρ	0.0003 (0.36)	-0.0008 (-0.40)	-2.84×10^{-5} (-0.07)	9.12×10^{-5} (0.09)	-0.0002 (-0.08)	0.0001 (0.25)

HDD period regroups January, February, March, April, October, November, December, CDD period spans from May to September, n denotes the number of observations and the figure in parentheses refers to the t-statistic.

7. Results

We want here to compute the monthly New York CDD weather futures prices for the period from 06/01/2006 to 06/30/2006 with the use of the estimates of the risk aversion parameter by the GMM and SMM approaches. The estimation of the coefficient by the GMM proceeding is carried out with the weather futures quotations for the period from 04/01/2004 to 12/31/2005 which represents 443 observations and for the CDD period covering from 05/01/2005 to 09/30/2005 which consists of 107 prices. We need for this method to define the instrumental variables. Since we have little guidance in choosing their number, we use three different sets of instruments. In the first one, we include the constant and two lagged values of the consumption growth rate. In the second one, we regroup the constant and one lagged value of the consumption growth rate and return. In the third set, we form the vector of instruments by using the constant and two lagged values of the consumption growth rate and return. Ferson and Constantinides (1991) point out that other variables than these lagged values should be employed as instruments to avoid the measurement errors to bias the parameter

estimates as well as the test of overidentifying restrictions. It is difficult in our case to find variables that are good instruments, i.e., that are highly correlated with both the consumption and the return of the weather futures. For instance, the personal disposable income is very correlated with the consumption expenditures but not with the weather futures prices. The GDP is surely correlated with the two variables but its frequency is a year or a quarter. So we cannot employ it to calculate the daily prices. Therefore, we limit our study to the first three proposed sets which represent respectively 3, 3 and 5 moment conditions. The results of the GMM principle are shown in Table 9 and Table 10. For the estimation of the coefficients by the SMM approach, we use a shorter period for the quotations which is from 06/01/2005 to 06/30/2005 which corresponds to 30 moment restrictions (in fact 22 since the quotations are 5 day week). With the SMM procedure, instead of the number of instruments, we need to determine the number of simulations N to run. We choose $N=2000$ for our experimentations. The findings are presented in Table 11.

We can notice in Table 9 that the estimated value of the risk aversion coefficient from the GMM approach is not stable. It varies with the set of instruments and with the bandwidth chosen for the estimation. Moreover in some cases, it is negative. We see that these values are particularly biased when the sample size is small (Table 10). The test of overidentifying restrictions indicates that the model is well specified in all the cases even when the risk aversion coefficient has the wrong sign. Indeed, the probability that the J-statistic exceeds the calculated value is always well above the 5% level. Lund and Engsted (1996) put also in evidence these aberrations. The fact that the test is biased in favor of accepting the model may be due to the use of the lagged values of the consumption growth rate and return. Another point which is also mentioned in some studies and that we can remark in Table 9 and Table 10 is that the prewhitening method enables to improve the confidence interval coverage but at the cost of inflating the value and the variance of the estimators. In some settings, it increases the bias as well as the nonconvergence problems especially for the small sample sizes (see also Andersen and Sørensen (1996) and Chiang and Kao (2005)). When we compare the results provided by the GMM principle to those of the SMM approach in Table 11, we see that they are quite different. The hypothesis that the model fits the data is rejected. The estimates of the risk aversion coefficient are greater than those presented in Table 9 and Table 10 (without the prewhitening procedure) because the simulated consumption is smoother than the observed one which induces a smaller value of the consumption ratio in the case of the SMM approach than for the GMM proceeding. So a more important size of the risk aversion parameter is needed to diminish the level of the estimated prices to match the quotations. We illustrate in Figure 3 the weather futures prices obtained when applying the SMM findings and in Figure 4 the results when employing the GMM estimates. For the last case, we choose the values resulted from the set 3 of instrumental variables and the Andrews (1991)'s bandwidth selection with prewhitening since they are the nearest to those of the SMM proceeding. We observe that the calculated prices fit better the quotations when the SMM estimates are used. To allow the computed prices to match the observations, the derived risk aversion parameter must be very high whereas the range of its plausible values is between 0 and 10 (see Mehra (2003)). As we exposed in section 2, numerous studies pointed out this failure of the consumption-based asset pricing model and attempted to bring some explanations. Among them, the Mankiw and Zeldes (1991)'s and Constantinides and Duffie (1996)'s points of view are particularly interesting. As we mentioned before, Mankiw and Zeldes (1991) showed that it was very high because the total consumption which was commonly used for the estimation was less correlated with the asset movements than the stockholder consumption. Constantinides and Duffie (1996) argued that the estimated coefficient did not correspond to the value of the risk aversion parameter. It is in fact a

function of it when individuals have heterogeneous income. They demonstrate that, in the presence of shocks to the labor income of each agent, the Euler condition includes in addition the cross-sectional variance of the individual consumers' consumption growth. This Euler condition can then be turned into the familiar form by assuming that the shocks are persistent, independently and identically distributed. The resulted coefficient that appears in this last representation is not the risk aversion coefficient but is a function of it. We can attempt to derive the "true" value of the risk aversion parameter by applying the Constantinides and Duffie (1996)'s formula and to see whether it belongs to the range of realistic values but this undertaking is beyond the objective of this paper which is to find the estimates of the parameters of the Euler condition so that the calculated weather futures prices reproduce correctly the observed quotations.

Table 9 : Estimation of the coefficients of the Euler conditions by the GMM procedure for the period from 04/01/2004 to 12/31/2005

	Set 1 (n=443 and L=3)		Set 2 (n=443 and L=3)		Set 3 (n=443 and L=5)	
	B	γ	β	γ	β	γ
Newey and West (1994)'s bandwidth selection	0.998 (702.69)	-50.040 (-1.45)	1.000 (791.19)	12.885 (0.91)	0.999 (865.04)	3.157 (0.25)
b	15		15		15	
J-statistic	1.58		3.38		4.68	
Probability	0.21		0.06		0.20	
Newey and West (1994)'s bandwidth selection + prewhitening	0.998 (172.07)	-54.501 (-0.50)	1.003 (225.23)	71.548 (1.04)	1.003 (233.64)	68.04 (1.03)
b	0		8		7	
J-statistic	0.007		0.73		1.28	
Probability	0.93		0.39		0.73	
Andrews (1991)'s bandwidth selection	0.998 (665.44)	-53.758 (-1.57)	0.999 (752.06)	11.795 (0.75)	0.999 (881.38)	0.230 (0.02)
b	31		37		36	
J-statistic	1.58		2.77		3.75	
Probability	0.21		0.09		0.29	
Andrews (1991)'s bandwidth selection + prewhitening	0.998 (149.02)	-57.561 (-0.47)	1.004 (188.12)	77.761 (0.88)	1.003 (202)	67.341 (0.82)
b	2		2		2	
J-statistic	0.0118		0.89		1.15	
Probability	0.91		0.34		0.76	

The New York weather futures quotations and the US real per-capita consumption are used for the period from 04/01/2004 to 12/31/2005. n is the number of observations, b denotes the bandwidth and the figure in parentheses refers to the t-statistic. "Probability" corresponds to the probability that the J-statistic with L-p degrees of freedom exceeds the calculated value where L and p are respectively the number of conditions and unknowns.

Table 10 : Estimation of the coefficients of the Euler conditions by the GMM procedure for the period from 10/01/2005 to 04/30/2005

	Set 1 (n=107 and L=3)		Set 2 (n=107 and L=3)		Set 3 (n=107 and L=5)	
	B	γ	β	γ	β	γ
Newey and West (1994)'s bandwidth selection	0.925 (45.82)	-33.224 (-1.40)	0.915 (54.31)	-31.554 (-1.28)	0.940 (65.03)	-7.815 (-0.53)
b	8		8		8	
J-statistic	2.10		2.15		3.52	
Probability	0.15		0.14		0.32	
Newey and West (1994)'s bandwidth selection + prewhitening	0.852 (15.12)	-221.462 (-0.62)				
b	7		No convergence		No convergence	
J-statistic	0.03					
Probability	0.87					
Andrews (1991)'s bandwidth selection	0.923 (82.99)	-40.836 (-1.57)	0.920 (91.60)	-36.428 (-1.92)	0.947 (107.80)	-23.281 (-1.70)
B	25		27		25	
J-statistic	1.55		1.53		3.63	
Probability	0.21		0.22		0.30	
Andrews (1991)'s bandwidth selection + prewhitening	0.870 (20.75)	-199.024 (-0.53)	0.914 (49.85)	-36.866 (-0.17)	0.921 (65.80)	-27.693 (-0.16)
b	2		0		0	
J-statistic	0.01		0.45		0.48	
Probability	0.91		0.50		0.92	

The New York weather futures quotations and the US real per-capita consumption are used for the period from 10/01/2005 to 04/30/2005 (CDD period). n is the number of observations, b denotes the bandwidth and the figure in parentheses refers to the t-statistic. "Probability" corresponds to the probability that the J-statistic with L-p degrees of freedom exceeds the calculated value where L and p are respectively the number of conditions and unknowns.

Table 11 : Estimation of the coefficients of the Euler conditions by the SMM procedure for the period from 06/01/2005 to 06/30/2005.

N=2000 and L=22		
	β	γ
Newey and West (1994)'s bandwidth selection	0.999 (471.53)	61.736 (1.76)
b	8	
J-statistic	221.69	
Probability	0.00	
Andrews (1991)'s bandwidth selection	0.999 (515.63)	61.706 (1.91)
b	24	
J-statistic	80.23	
Probability	0.00	

N is the number of simulations for the temperature and the consumption, b denotes the bandwidth and the figure in parentheses refers to the t-statistic. "Probability" corresponds to the probability that the J-statistic with L-p degrees of freedom exceeds the calculated value where L and p are respectively the number of conditions and unknowns.

Figure 3 : New York CDD weather futures prices (for working days) expiring in June 2006. Prices are calculated with the estimates of the SMM approach which was implemented with the Andrews (1991)'s bandwidth selection without prewhitening and by using 2000 simulations.

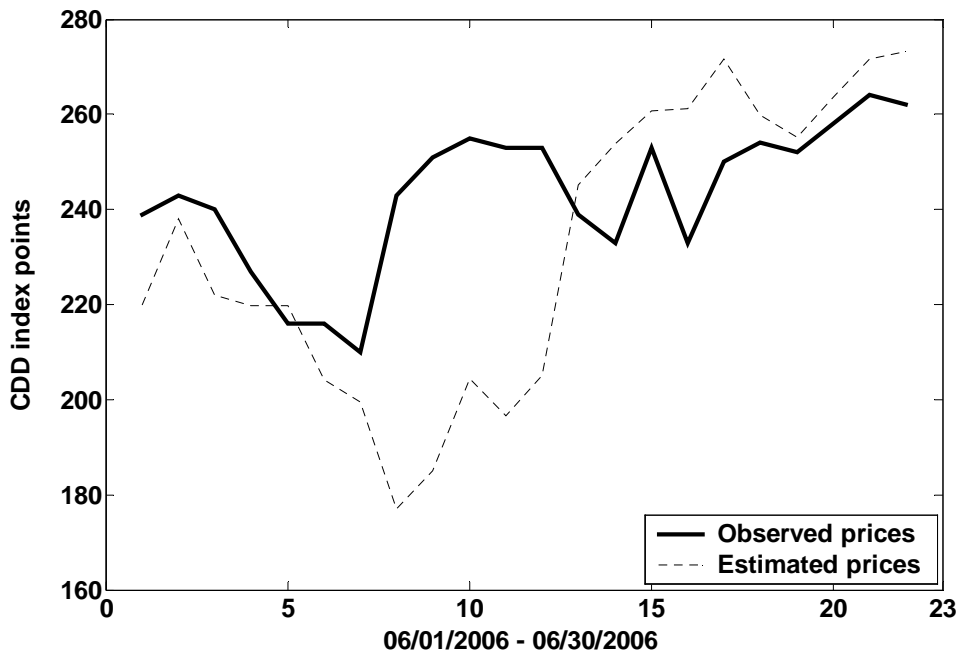
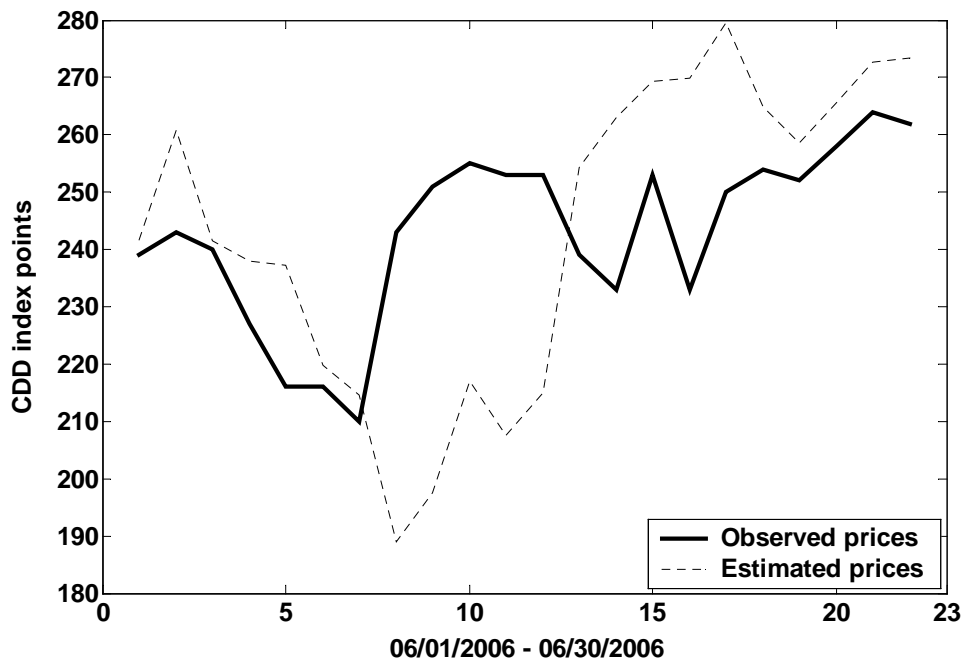


Figure 4 : New York CDD weather futures prices (for working days) expiring in June 2006. Prices are calculated with the estimates of the GMM approach which was implemented with the set 3 of instruments and the Andrews (1991)'s bandwidth selection with prewhitening and by using 2000 simulations.



8. Conclusion

In this paper, we have computed the New York weather futures prices by using the consumption-based asset pricing model. The parameters of this model were estimated by applying two methods which are the GMM and SMM approaches. The last one was implemented by simulating the New York daily average temperature and the per-capita consumption growth rate processes. We have shown that for the modelling of the temperature behavior the process composed of the Cao and Wei (2004)'s suggestion for the seasonality in the mean and the Benth and Šaltytė-Benth (2005)'s proposal for the seasonality in the variance is the more appropriate. The calculated prices resulting from the use of the SMM estimates have appeared to be the nearest to the observations. We plan in our future research to analyse the consequences of the use of the sectorial indexes (for instance in energy) instead of the consumption to compute the weather derivatives prices. The advantage of this approach is that no interpolation is needed to convert the data into a daily series since the frequency of the sectorial indexes is a day. Furthermore, this substitute can be more correlated with the temperature than the consumption.

Footnotes

(1) Cao and Wei (2004) take negative values for the risk aversion parameter because their utility function is of the form : $U(C_t) = \frac{C_t^{1+\gamma}}{1+\gamma}$.

(2) Asset return and consumption growth rate are regarded as stationary variables.

(3) When the errors $g_t(\hat{\theta})$ are uncorrelated, \hat{S} is written as $\hat{S} = \frac{1}{n} \sum_{t=1}^n g_t(\hat{\theta}) g_t(\hat{\theta})'$.

(4) Campbell and Diebold (2004) omit the constant term. The problem here is that the variance can be negative since its cyclical pattern has positive and negative phases.

(5) The Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) information criteria are respectively expressed as :

$$AIC = -2 \left(\frac{LL}{n} \right) + \frac{2k}{n} ,$$

$$SC = -2 \left(\frac{LL}{n} \right) + k \frac{\ln(n)}{n} ,$$

and

$$HQ = -2 \left(\frac{LL}{n} \right) + 2k \frac{\ln(\ln(n))}{n}$$

where LL corresponds to the log-likelihood of the model with k parameters.

(6) The H statistic of Durbin is given by :

$$H = \left(1 - \frac{DW}{2}\right) \sqrt{\frac{n}{1 - n \hat{\sigma}_{AR(1)}^2}}$$

where DW refers to the estimated value of the Durbin-Watson statistic, $\hat{\sigma}_{AR(1)}^2$ denotes the estimated variance of the first coefficient of the AR(5) process and n corresponds to the number of observations. The H statistic follows asymptotically a centered normal distribution with a variance one. When the term under the square root is negative, we cannot compute the statistic. In that case, autocorrelation of the errors is assumed.

References

- Andersen T.G. and Sørensen B.E. (1996), "GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study", *Journal of Business & Economic Statistics*, Vol.14, No.3, 328-352.
- Andrews D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica*, Vol.59, No.3, 817-858.
- Andrews D.W.K. and Monahan J.C. (1992), "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator", *Econometrica*, Vol.60, No.4, 953-966.
- Augros J.C. and Moréno M. (2002), *Les Dérivés Financiers et d'Assurance*, Economica.
- Benth F.E. (2003), "On Arbitrage-free Pricing of Weather Derivatives based on Fractional Brownian Motion", *Applied Mathematical Finance*, Vol.10, 303-324.
- Benth F.E. and Šaltytė-Benth J. (2005), "Stochastic Modelling of Temperature Variations with a View Towards Weather Derivatives", *Applied Mathematical Finance*, Vol.12, No.1, 53-85.
- Brix A., Jewson S. and Ziehmann C. (2002), "Weather Derivative Modelling and Valuation : a Statistical Perspective", in *Climate Risk and the Weather Market, Financial Risk Management with Weather Hedges*, ed. Dischel R.S., 127-150.
- Brody D.C., Syroka J. And Zervos M. (2002), "Dynamical Pricing of Weather Derivatives", *Quantitative Finance*, Vol.2, 189-198.
- Caballero R., Jewson S. and Brix A. (2002), "Long Memory in Surface Air Temperature: Detection, Modelling and Application to Weather Derivative Valuation", *Climate Research*, Vol. 21, 127-140.
- Campbell S.D. and Diebold F.X. (2004), "Weather Forecasting for Weather Derivatives", CFS Working Paper.
- Cao M. and Wei J. (2004), "Weather Derivatives Valuation and Market Price of Weather Risk", *The Journal of Futures Markets*, Vol.24, No.11, 1065-1089.
- Chiang M. and Kao C. (2005), "Spectral Density Bandwidth Choice and Prewhitening in the Generalized Method of Moments Estimators for the Asset Pricing Model", *Economics Bulletin*, Vol.3, No.10, 1-13.
- Constantinides G.M. and Duffie D. (1996), "Asset Pricing with Heterogeneous Consumers", *Journal of Political Economy*, Vol.104, No.2, 219-240.
- Duffie D. and Singleton K.J. (1993), "Simulated Moments Estimation of Markov Models of Asset Prices", *Econometrica*, Vol.61, No.4, 929-952.

- Ferson W.E. and Constantinides G.M. (1991), “Habit Persistence and Durability in Aggregate Consumption”, *Journal of Financial Economics*, Vol.29, 199-240.
- Hamisultane H. (2006a), “Extracting Information from the Market to Price the Weather Derivatives”, Working paper.
- Hamisultane H. (2006b), “Pricing the Weather Derivatives in the Presence of Long Memory in Temperatures”, Working paper.
- Hansen L.P. (1982), “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, Vol.50, No.4, 1029-1054.
- Hansen L.P. and Singleton K.J. (1982), “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models”, *Econometrica*, Vol.50, No.5, 1269-1286.
- Jewson S. (2004), “Introduction to Weather Derivative Pricing”, Risk Management Solutions Inc.
- Lee B. and Ingram B.F. (1991), “Simulation Estimation of Time-Series Models”, *Journal of Econometrics*, Vol.47, 197-205.
- Lucas R.E. Jr. (1978), “Asset Prices in an Exchange Economy”, *Econometrica*, Vol.46, No.6, 1429-1445.
- Lund J. and Engsted T. (1996), “GMM and Present Value Tests of the C-CAPM: Evidence from the Danish, German, Swedish and UK Stock Markets”, *Journal of International Money and Finance*, Vol.15, No.4, 497-521.
- Mankiw N.G. and Zeldes S.P. (1991), “The Consumption of Stockholders and Nonstockholders”, *Journal of Financial Economics*, Vol.29, 97-112.
- Mehra R. (2003), “The Equity Premium: Why Is It a Puzzle?”, *Financial Analysts Journal*, 54-69.
- Mehra R. and Prescott E.C. (1985), “The Equity Premium: a Puzzle”, *Journal of Monetary Economics*, Vol.15, 145-161.
- Moréno M. (2003), “Weather Derivatives Hedging and Swap Illiquidity”, Weather Risk Management Association.
- Newey W.K. and West K.D. (1987), “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica*, Vol.55, No.3, 703-708.
- Newey W.K. and West K.D. (1994), “Automatic Lag Selection in Covariance Matrix Estimation”, *Review of Economic Studies*, Vol.61, 631-653.
- Platen E. and West J. (2004), “Fair Pricing of Weather Derivatives”, Working paper.
- Richards T.J., Manfredo M.R. and Sanders D.R. (2004), “Pricing Weather Derivatives”, *American Journal of Agricultural Economics*, Vol.4, 1005-1017.

Roustant O. (2002), “Une Application de Deux Modèles Econométriques de Température à la Gestion des Risques Climatiques (1^{ère} partie)”, *Banque & Marchés*, 58, 22-29.

Roustant O. (2002), “Une Application de Deux Modèles Econométriques de Température à la Gestion des Risques Climatiques (2^{ème} partie)”, *Banque & Marchés*, 59, 36-44.

Weil P. (1989), “The Equity Premium Puzzle and the Risk-Free Rate Puzzle”, *Journal of Monetary Economics*, Vol.24, No.3, 401-421.