

## *Riding the Temp*

Michael Moreno  
ISFA - Universite Lyon 1  
43, bd du 11 novembre 1918  
69622 - Villeurbanne CEDEX

in association with  
Speedwell Weather Derivatives\*  
19, St Mary At Hill  
London EC3R 8EE  
44. (0)207.929.79.79  
<http://www.weatherderivs.com/>

### **Abstract :**

Two processes have been extensively discussed for the simulation of the temperature in order to price weather options. Past studies revealed that these kinds of processes were fitting correctly the temperature data. However no direct comparison of these two processes have been written yet. Firstly explaining these two processes, we then look at their goodness of fit for Marseille and Paris-Orly. Going further than past studies, the third part of this article finally reveals if one can use such processes to simulate the temperature.

Keywords : Option, Temperature, Weather.

---

\* Speedwell Derivatives Limited is regulated by the SFA.

**T**emperature based contracts are currently the most traded in the weather market. As an example, the EnronOnline weather trading system [6] offers no less than 99 weather contracts that can be traded real-time. All are based on Heating Degree-Day (HDD) or Cooling Degree-Day (CDD). HDD and CDD are cumulative differences between a reference temperature (65°F) and the average of the daily high and low. The current dominance of HDD and CDD contracts arose from the protection against weather needed by energy companies. Nevertheless, irrespective of the required protection (see Life Outside the energy industry [5]) it is worthwhile constructing a simulation of the temperature process instead of statistically estimating prices using HDD data. Indeed, many problems can be encountered with straightforward estimation of HDD or CDD distributions. The first one is that irrespective of the extent to which the average temperature on a given day is below the reference, such a day contributes zero to the cumulative index. But clearly the nature of the risk in the underlying temperature process is different depending on whether that daily average was, say, 64.5 or 0. Secondly using the CDD or HDD series directly is likely to mean that only 30 or 40 points are available. Moreover, suppose that only 10 years of data are available or that there is a strong case that only the last ten years are representative of the likely future distribution of the index: it may then be hard to estimate the distribution with a good fit. More problems come to light when trying to estimate a contract that is currently within its reference period. Using a statistical approach will inevitably drive one to an estimation based on conditional probability. Using a temperature simulation is likely to be a superior method. As a consequence much effort on the part of banks, insurance companies and energy companies has been expended on simulating the temperature process.

While temperature is often continuously measured, the values used in order to calculate HDD or CDD are the daily average of the temperature. These values are therefore discrete. Using a continuous process to estimate the daily average temperature process in order to forecast it is consequently not necessary. Therefore instead of starting with a continuous mean-reverting model as did Dischel [3] or Dornier & Queruel (DQ hereafter) [4] and then to discrete it, discrete processes could be used directly. The main purpose of this article is to compare the two discrete processes for modelling the temperature process that have been published in previous papers. The first one is the well known mean reverting process that we explain building it in discrete time. The second one is that introduced by Carmona [2]. We will finally conclude on their goodness of fit.

## **I. Presentation of the two processes**

### **A. The mean reverting process**

The daily average temperature  $T_1$  can be viewed as a time series process. Figure 1 visualises this.

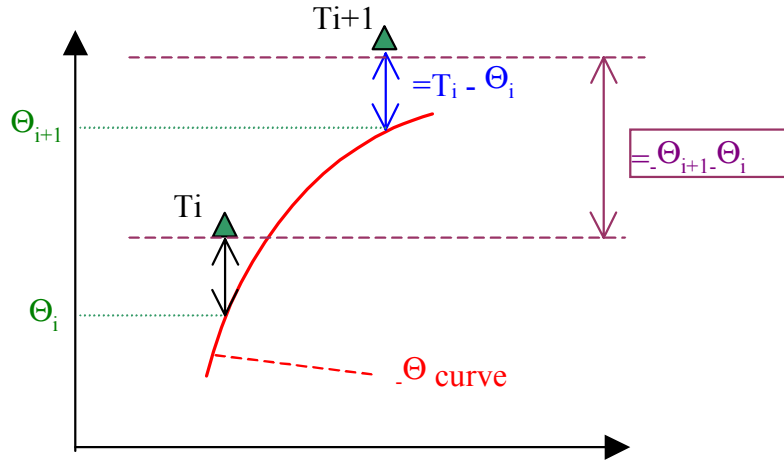


figure 1 : Building first property of  $T_i$ .

The temperature  $T_{i+1}$  at time  $i+1$  can be built starting with the following assumptions:

- It depends on the temperature at time  $T_i$
- It depends on the increase or the decrease of the average temperature between date  $i$  and  $i+1$

So, the actual “deterministic process” can be written as:

$$T_{i+1} = T_i + (\Theta_{i+1} - \Theta_i) \quad (1)$$

Where  $\Theta_i$  is the mean of the temperature at date  $i$ .

The actual “process” is completely deterministic. In order to create some uncertainty, a noise  $\phi_i$  is added for each date  $i$  (see figure 2).

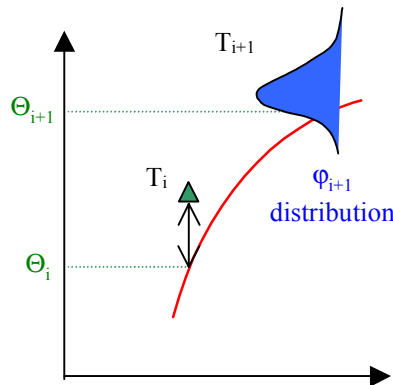


figure 2 : Adding the uncertainty to  $T_{i+1}$ .

So we rewrite the process as :

$$T_{i+1} = T_i + (\Theta_{i+1} - \Theta_i) + \phi_i \quad (2)$$

No conditions are made on the noise at this stage.

But, one of the most well known properties of temperature is that it is mean reverting to the global average temperature. Expressed otherwise, the noise  $\phi_i$  is conditional on the past and should be written as:

$$\phi_i = \alpha(\Theta_{i-1} - T_{i-1}) + \varepsilon_i \quad (3)$$

Where  $\alpha$  is a constant and  $\varepsilon_i$  is a “complementary” noise.

To end this first part of the construction of the process, we now have:

$$T_{i+1} = T_i + (\Theta_{i+1} - \Theta_i) + \alpha(\Theta_i - T_i) + \varepsilon_i \quad (4)$$

This process generalises the discretisation of the extensively described continuous process:  
 $dT_t = \alpha(\Theta_t - T_t)dt + \gamma_t dW_t$

Where  $W_t$  is a Wiener process and  $\gamma_t$  a deterministic function of time  $t$ .

Indeed, if we restrict  $\varepsilon_i$  to the following process:  $\varepsilon_i = \gamma_i \text{GWN}_i$  with  $\text{GWN}_i$  a Gaussian white noise then we obtain the corrected process developed by DQ which was generated from the discretisation of the continuous process. DQ asserted that without losing any generality one could suppose  $\gamma_i$  constant. Unfortunately, this last hypothesis is not always adapted to represent the real temperature process.

## B. The Autoregressive Process

The process presented above is based on a difference-process. This can be show since equation 2 could be rewritten:

$$(T_{i+1} - T_i) - (\Theta_{i+1} - \Theta_i) = \phi_i \quad (5)$$

Carmona proposed instead to simulate the process using the following assumption:

$$T_{i+1} = \Theta_{i+1} + AR(p) \quad (6)$$

The memory of the process of the temperature appears more “clearly” through the use of the autoregressive process with  $p$ -order  $AR(p)$ .

Now if we rearrange equations (5) and (6) then we obtain:

$$T_{i+1} - \Theta_{i+1} = (T_i - \Theta_i) + \phi_i \quad (5')$$

and

$$T_{i+1} - \Theta_{i+1} = AR(p) \quad (6')$$

Thus, if both processes fit the data then the  $AR(p)$  process should be as follows:  $AR(p) = (T_i - \Theta_i) + \phi_i$ .

## II. Estimations and numerical comparisons of the process

### A. Estimating the mean-reverting process

Suppose that the temperature process is given by (4). The estimation of the noise  $\varepsilon_i$  is described in the following. The data used are Paris-Orly daily average temperatures. The calculation of  $\Theta_i$  was determined with detrended temperature using 30 years of data then evaluated with a partial average. A sinusoid was deliberately avoided: the slope is too regular. It does not fit the asymmetric evolution of temperature in summer compared with winter. Nor does the increase in springtime temperatures usually reflect the decrease in autumn temperatures.

For the moment, supposing  $\gamma$  constant and without specifying any distributions of the white noise, the least squares method was used in order to obtain the value of  $\alpha$ . Then, the residues have been calculated. For Orly, the result residues  $\varepsilon_i$  against time are produced in figures 3, 4 & 5.

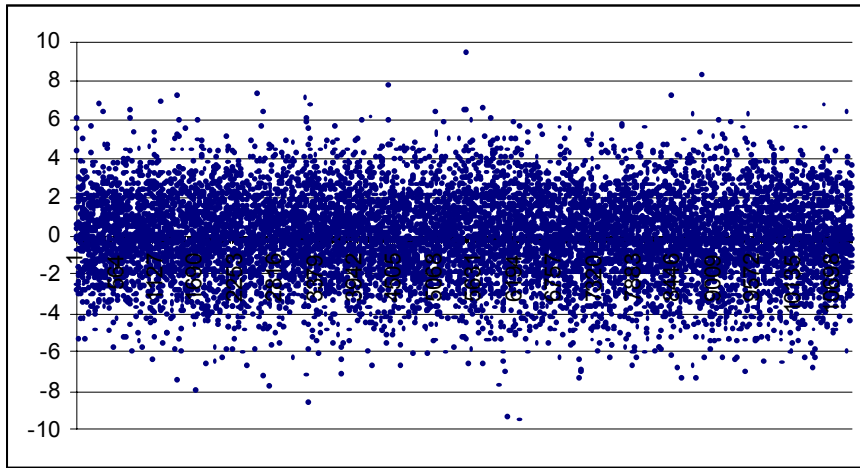


Figure 3: Paris – Orly Residues

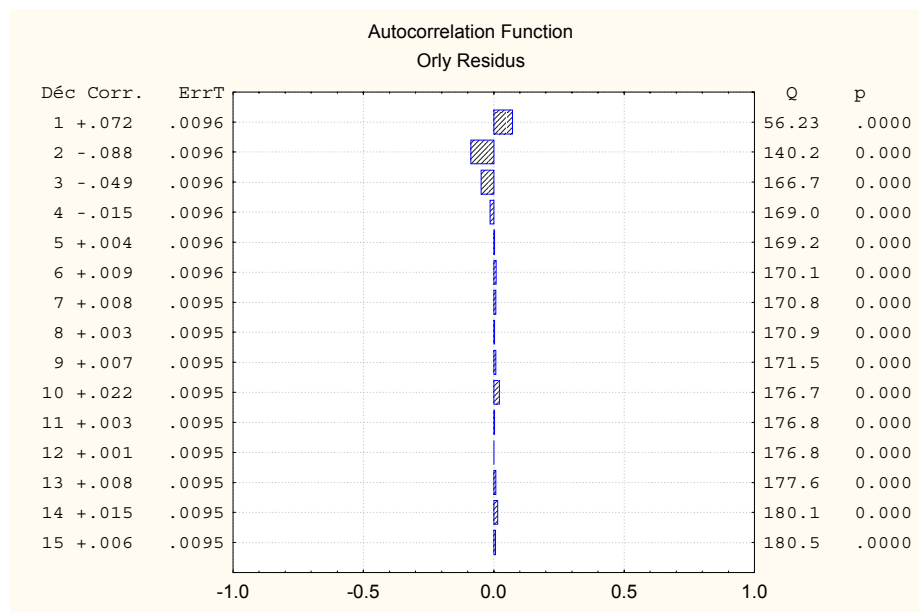


Figure 4 : Autocorrelation function of Paris-Orly residues

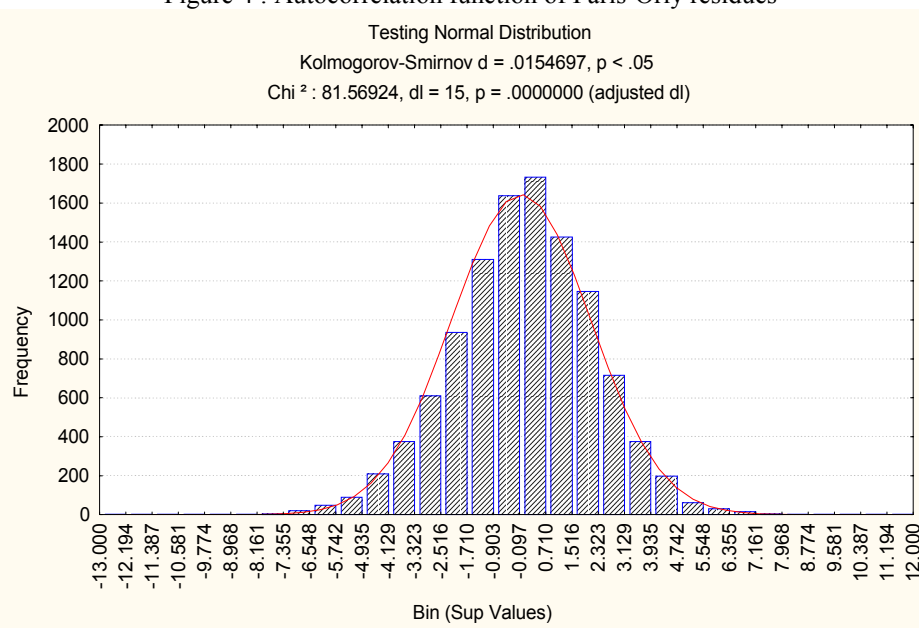


figure 5 : Overall distribution of Paris-Orly residues

Once could conclude from this point that the process fits quite well the data. However, given the bias visible in figure 5, the Gaussian white noise could be replaced by another noise with a slightly asymmetric distribution. (The Pearson distributions could be used for this). Repeating the above analysis with Marseille is a priori also conclusive (see figure 6):

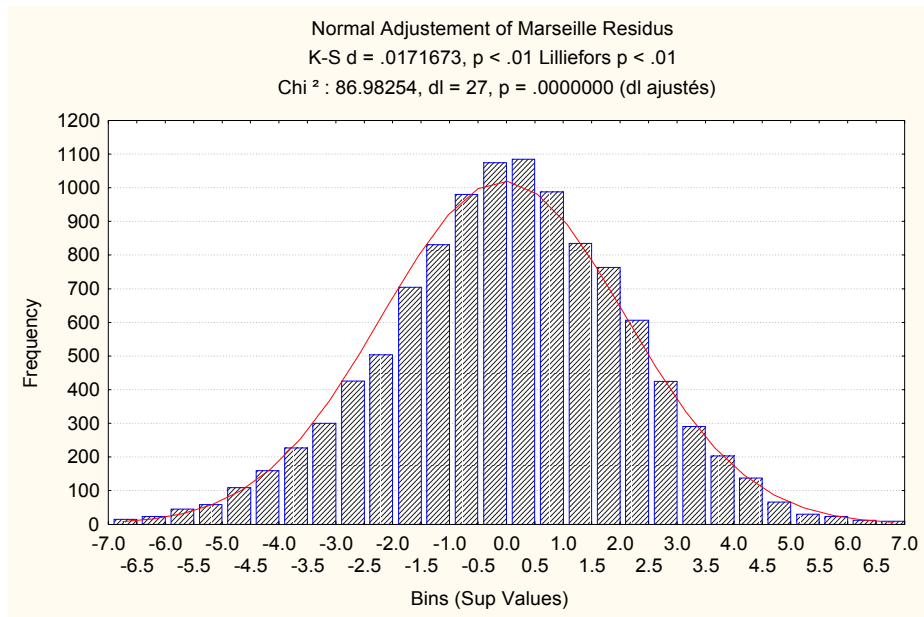


figure 6 : Normal adjustment of Marseille Residues

But, those tests are not complete. If we look at the distributions of the residues through time shall we still conclude the same goodness of fit. Figure 7 (resp. 8) shows the distribution of the volatility of the residues of Orly (resp. Marseille).

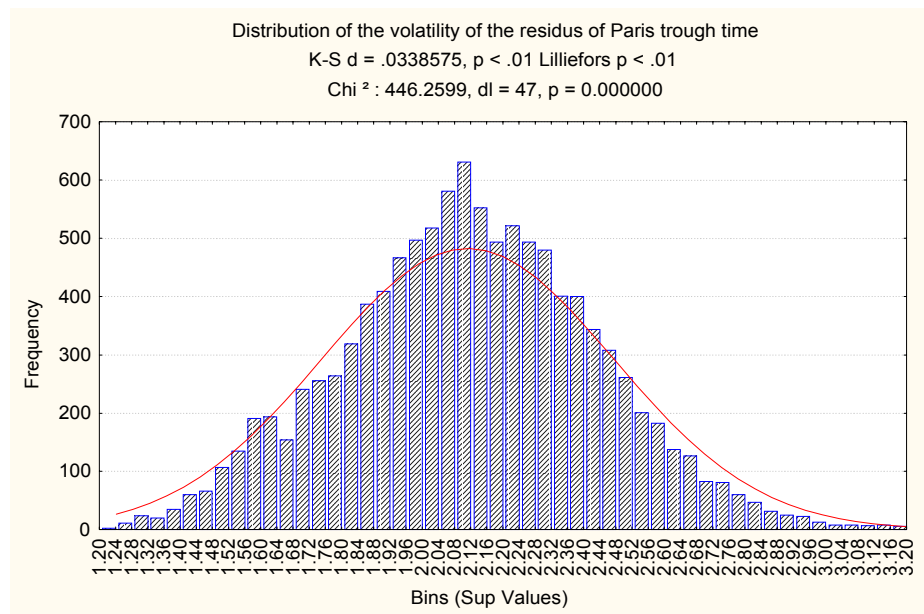


figure 7 :

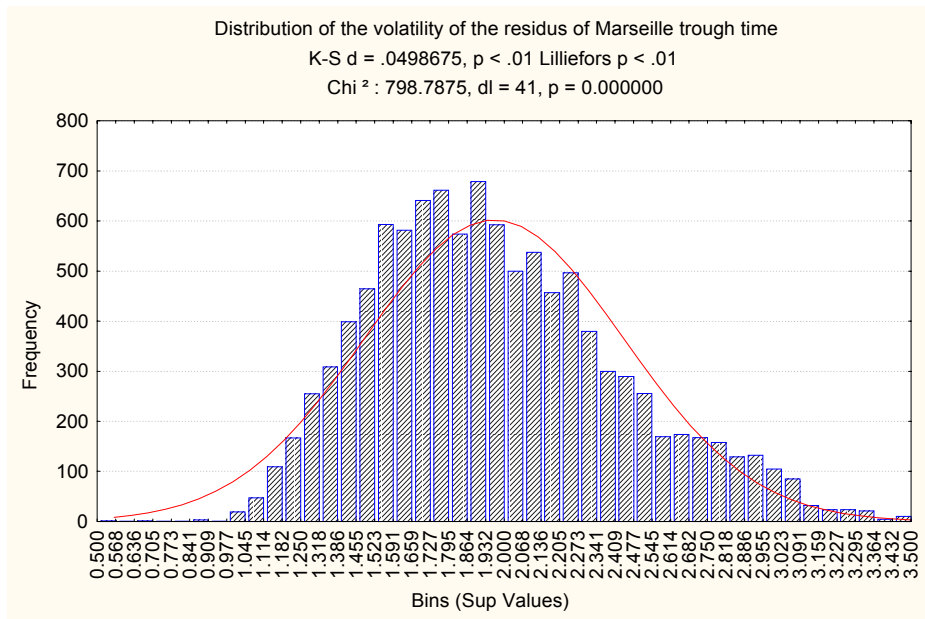


figure 8 :

Given the non-normal distribution of the volatility of the residues for Marseille further analysis is required. An investigation into possible seasonally effects is now carried out: Figure 9 shows the evolution of the 180 days moving average of the volatility of the residues calculated with a 30 days basis for both towns during 3650 days :

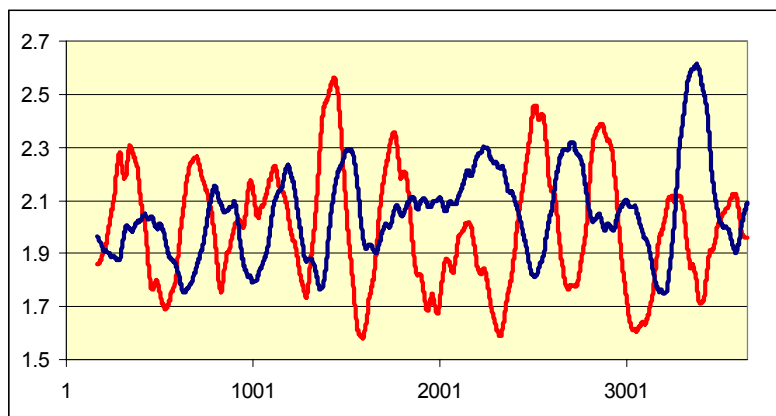


Figure 9 :

This last figure demonstrates that the volatility of the residues has a seasonal component that might have been included before. So presuming that  $\gamma$  is constant through time is incorrect. Volatility of temperature during the summer is different from that during winter. Analysis on both towns shows the volatility is higher in winter than in summer. In other words, the risk is different from winter to summer because the distribution of residues is non homogenous. Therefore process must be readjusted. The function of time  $\gamma(t)$  is chosen arbitrarily to be a sinusoid of the form :  $1 + \sin^2(\theta t + \beta)$ . The two above curves are then transformed to the following one :

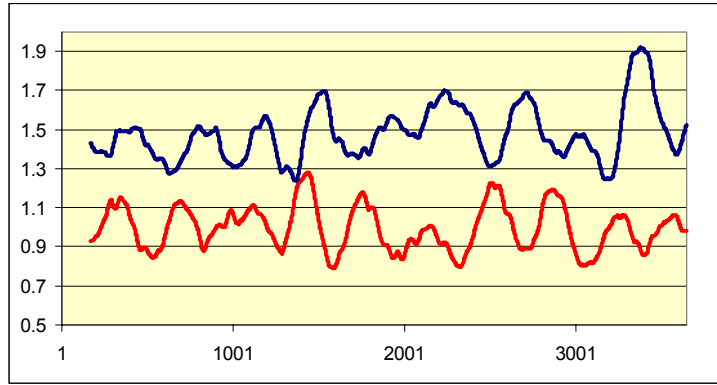


Figure 10 :

The curves show that including this new function for  $\gamma_t$  has considerably reduce the effect of the seasonally component the volatility still possesses some seasonnality. Using other sinusoid functions of volatility did not lead to significantly better fits.

## B. Estimating the AR process

We have seen that the temperature process appears adequately represented by a mean-reverting process with a white noise that is not perfectly Gaussian and with a seasonal volatility. We now turn to estimating the simple AR process in order to compare the results. In order to provide results we first represent the autocorrelation function of the difference:  $T_{i+1} - \Theta_{i+1}$ .

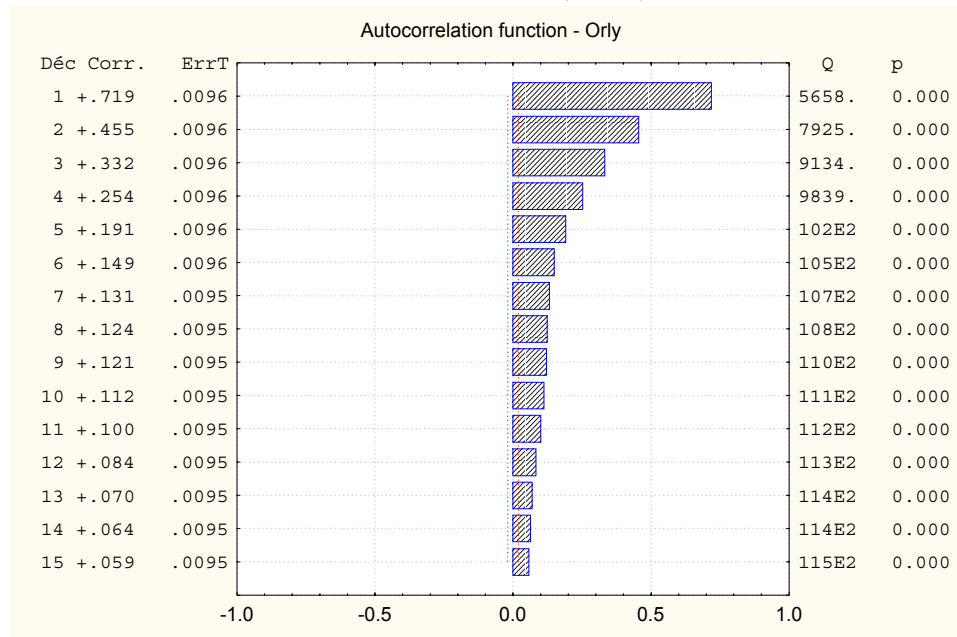


Figure 11:

With the information of the partial autocorrelation function and the one provided from this first study, the order of the AR process we have chosen to represent the temperature is 6. After having estimating the process we show that the new calculated residues are not correlated (Figure 12) and are normally distributed (Figure 13).



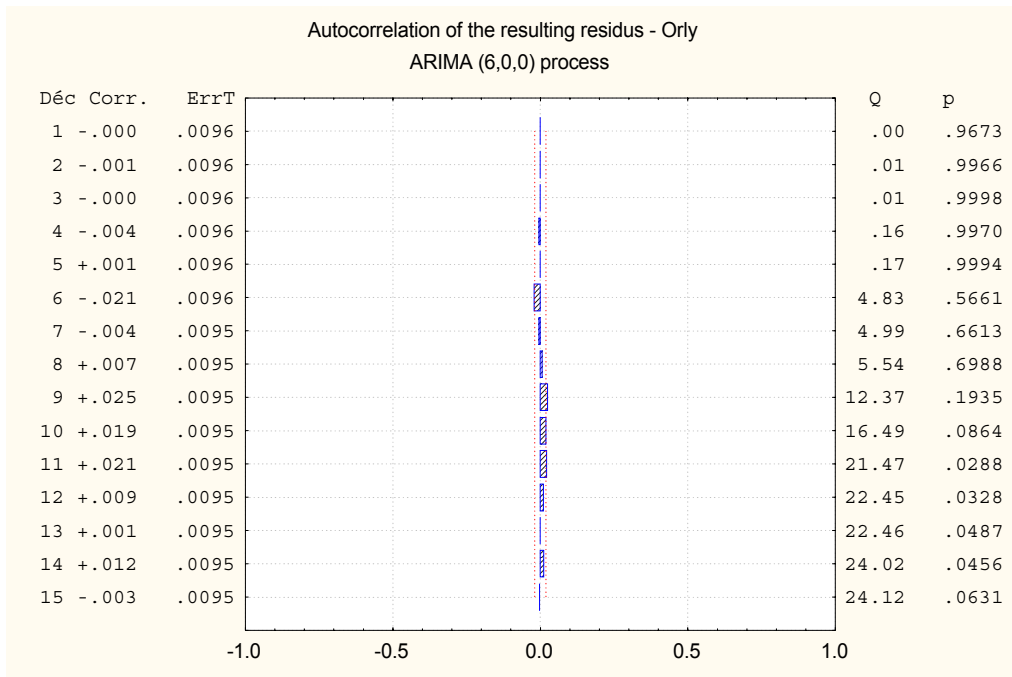


Figure 12:

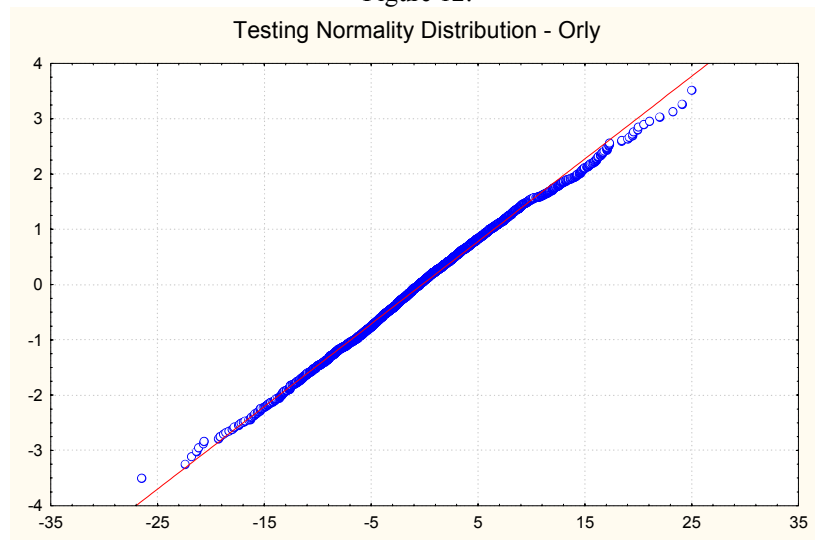


Figure 13:

Running the same estimation process with Marseille data we obtain similar results with a 7-order AR process (Figure 14, 15 & 16).

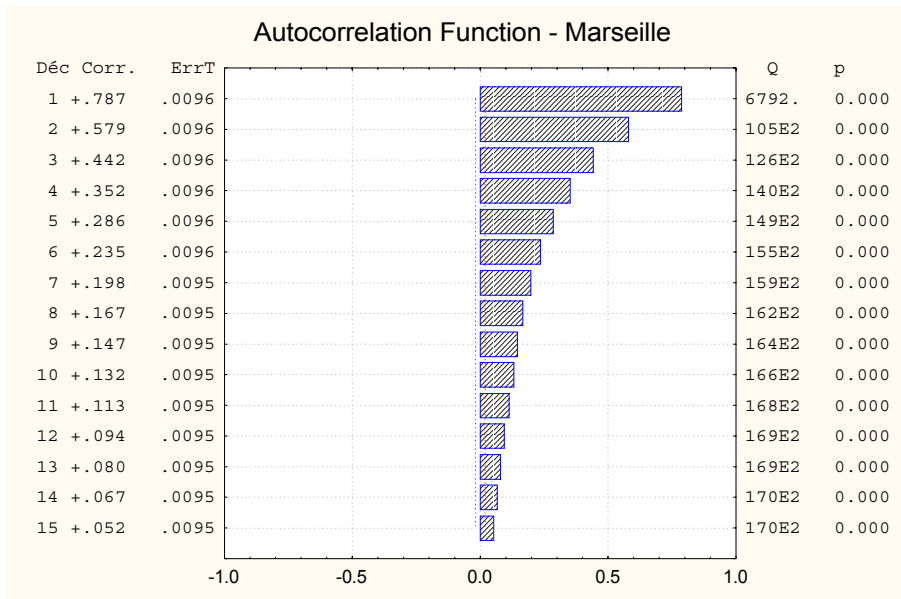


Figure 14:

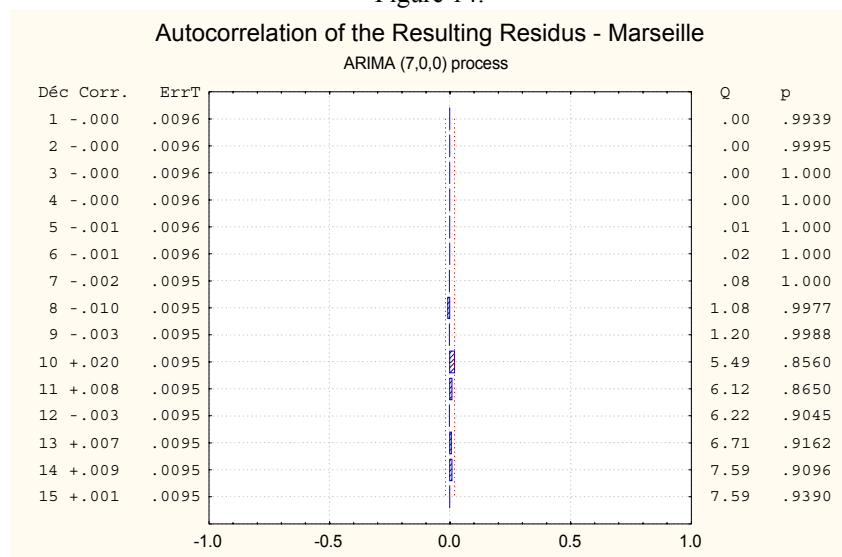


Figure 15

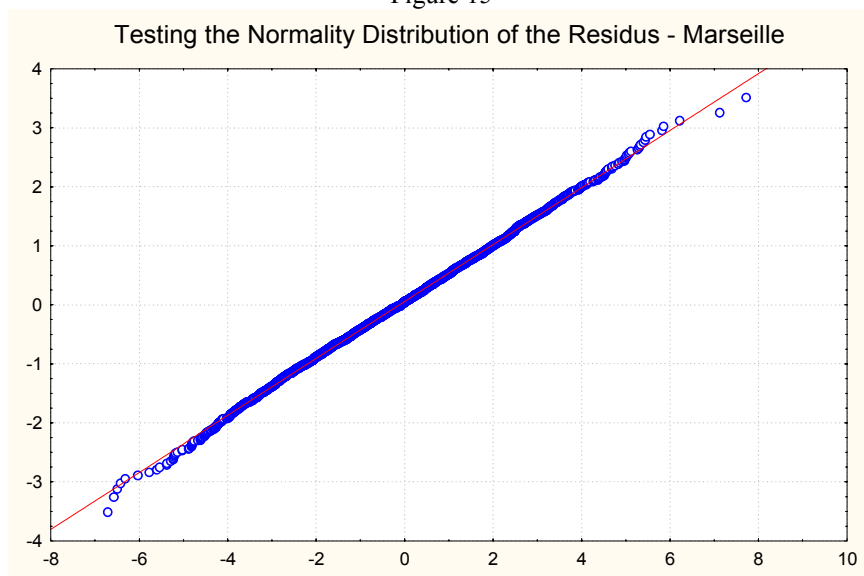
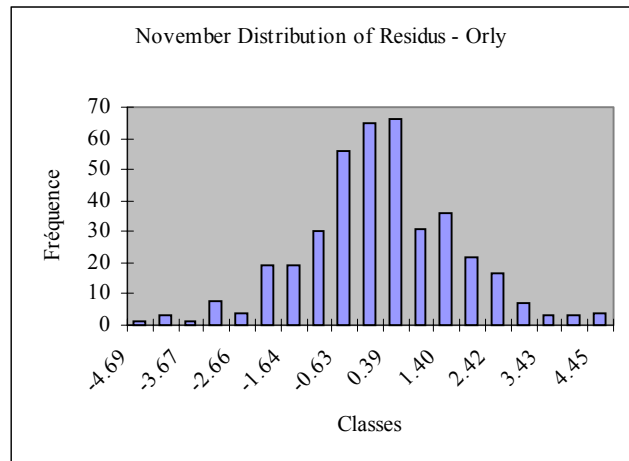
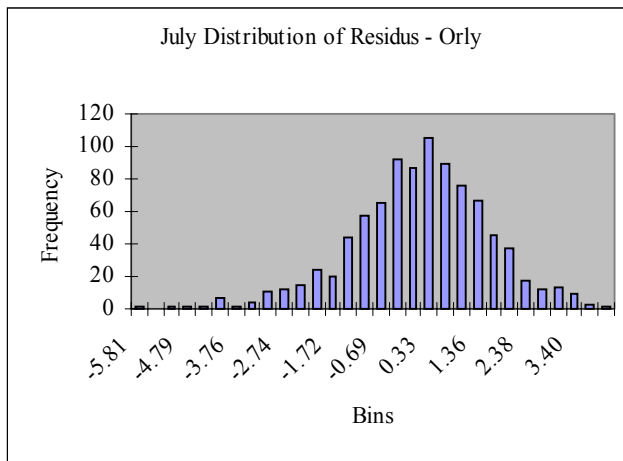
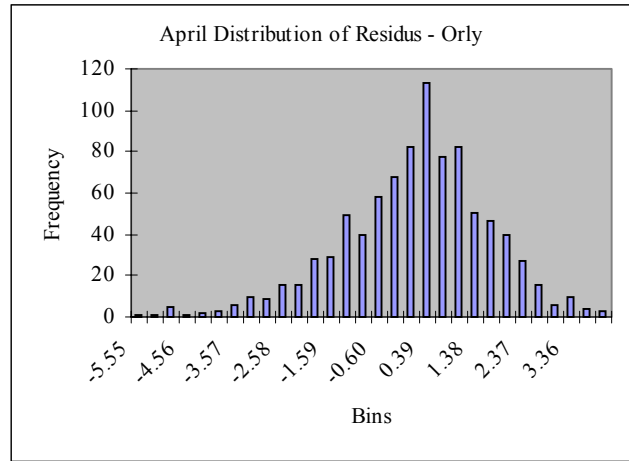
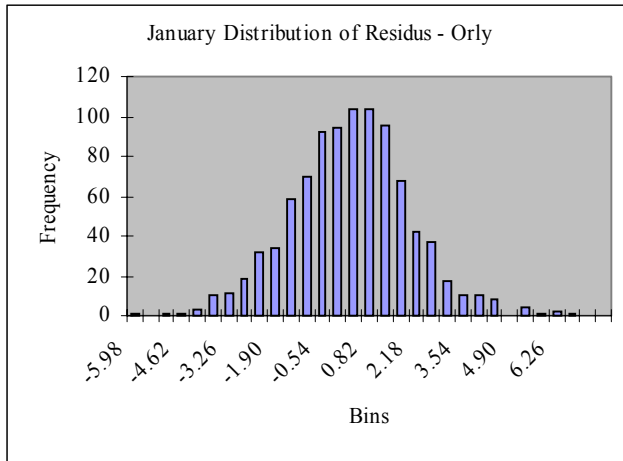


Figure 16

With these processes, the residues are normally distributed and the goodness of fit seems impressive and even better than with the mean reverting process.

### III. Are these processes really good?

Past studies have suggested that these processes fit the data well. It can be demonstrated with basis statistics that these processes may not be as good as previous studies have asserted. Indeed, if they were good then the conditional distributions of the residues given different periods of the year should be the same. The four following graphics show that it is not the case:



Clearly the distribution of the residues through the year is non-constant. Except the two first moments, the other moments differ from time to time. Hence the assumptions previously made to fit the process are not fully respected. The problem is not coming from the volatility of the residues. The problem is deeper: the final residues are not identically distributed, they are only independent. Whatever the town and the process we have estimated, non-homogenous distributions have been raised.

### Conclusion

Many papers suppose that the process of the temperature is either a mean-reverting process or an autoregressive process with a noise that is i.i.d. We have demonstrated that the goodness of fit appears satisfactory, but that these basic processes should not be used to simulate the temperature because the distribution of the noise is not homogenous through time. Other classes of processes should be employed.

## References

- [1] Brickwell P. & Davis R., "Time Series: Theory and Methods", Springer Series in Statistics, Second Edition, 1991.
- [2] Carmona R., "Calibrating Degree Days Options", Talk given at the Third Seminar on Stochastic Analysis, Random Fields and Applications, Ascona (Switzerland) September 23, 1999.
- [3] Dischel B., "Weather Derivatives", <http://www.adtrading.com/adt37/>.
- [4] Dornier F. & Queruel M., "Caution to the Wind", Weather Risk, August 2000.
- [5] Foster D. & Gautam J., "Come rain, come shine", Weather Risk, August 2000.
- [6] Foster K., "Optimistic Forecast", Weather Risk, August 2000.
- [7] Hamilton J., Time Series Analysis, Ed Princeton University Press, 1994.
- [8] Storch H. & Zwiers F., Statistical Analysis in Climate Research, Press Syndicate of the University of Cambridge, 1999.