Model risk in the pricing of weather derivatives

Olivier Roustant, Jean-Paul Laurent, Xavier Bay, Laurent Carraro

Model Risk in the Pricing of Weather Derivatives

Olivier Roustant (1,2), Jean-Paul Laurent (1), Xavier Bay (2), Laurent Carraro (2)

September 2003.

Abstract
Temperature modelling is a major issue for valuation of weather derivatives. Goodness of fit is usually assessed from historical data. However, estimation errors can result in large price uncertainty that may be problematic for practical applications. In this paper, we consider a temperature ARMA model and quantify the price uncertainties for weather Futures and weather options. Each price is seen as a *random variable* (which is a function of the parameters estimator), and we assess price uncertainty by giving confidence intervals. In addition, we look for sources of uncertainty, and point out the major defects of the model.

Résumé
La modélisation des températures est cruciale pour la valorisation des dérivés climatiques. Les paramètres des modèles sont estimés à partir de bases de données de températures. L'incertitude qui en résulte se répercute sur les évaluations des prix des dérivés climatiques. Dans cet article, nous quantifions ces effets pour des contrats futures et des options sur des indices de températures. Les prix des dérivés climatiques sont alors vus comme des variables aléatoires pour lesquelles nous donnons des intervalles de confiance. Nous déterminons également comment les différents paramètres du processus des températures contribuent à l'incertitude sur les prix des dérivés climatiques.

(1) Institut de Science Financière et d'Assurances, Université Claude Bernard (Lyon 1), 43 Boulevard du 11 novembre 1918, 69 622 Villeurbanne Cedex.
Tel : (33) (0)4.72.43.11.75. Fax : 04.72.43.11.76. e-mail : jean-paul.laurent@univ-lyon1.fr.
(2) Ecole Nationale Supérieure des Mines de Saint-Etienne, 158 Cours Fauriel, 42 023 Saint-Etienne Cedex 2.
Tel : (33) (0)4.77.42.66.75. Fax : 04.77.42.66.66. e-mail : roustant@emse.fr.

The authors thank F. Dufresne, C. Gouriéroux and O. Scaillet for useful comments. They also thank participants of the AFFI congress (June 2002, Strasbourg) for relevant remarks.
The data of temperature of the Paris-Montsouris station was provided by Météo-France.

Keywords: weather derivative, model risk, price uncertainty.

JEL: G13, C12, C15, C22.

I. INTRODUCTION

The market for weather derivatives was launched by investment banks, insurance companies and utilities in the late 90's. Most of the contracts are OTC though some can be traded on Future Exchanges (CME, LIFFE-Euronext). These products provide protection against losses due to non-catastrophic climatic events. End-users are, mostly, energy companies but also theme parks, breweries, winter shipment manufacturers, leisure resorts, fertiliser manufacturers… The underlying climatic risks are measured by means of indexes built from available meteorological data.

Since these markets are currently quite illiquid and not very transparent, it is difficult to mark to market the products and calibrate some parameters from market prices. Thus, market participants rather use econometric models plus a pricing rule and then mark to model. We want here to investigate the magnitude of model risk arising from this approach. We look for key parameters that are most influential on prices. The outcome of the paper is twofold: firstly, we show that parameters that drive the mean of the temperature are the most important, even for short-term products. This suggests that more emphasis should be put with respect to modelling temperature trends. Secondly, we can provide some confidence intervals for prices; this can be used for determining a reserve policy and better cope with model risk.

The article is organised as follows. In section II, we describe more precisely the model risk issue. In section III, we present the data, and some specifications for the weather derivatives used. In section IV, we quantify price uncertainty corresponding to (daily) temperature modelling. In particular, we look for sources of uncertainty by assessing specific influence of each parameter (or groups of parameters) on prices. Finally, some concluding remarks and directions for future researches are given in section V.

II. MODEL RISK METHODOLOGY

The researches on temperature modelling have led to some autoregressive econometric models, which are well suited for simulating temperatures over time horizons within some months, that is the usual maturity of weather contracts[1]. Earlier work relied upon a Hull and White AR(1) model (Dischel (1998), Dornier and Quéruel (2000) or Moréno (2000)). More recently, Brody, Syroka and Zervos (2001) have proposed to replace the traditional Brownian motion by a fractional Brownian motion, leading to a *fractional* Ornstein – Uhlenbeck, allowing to incorporate a long memory effect. Besides, Davis (2001) has used the geometric Brownian motion to model the accumulated HDDs (or CDDs). The second approach is based solely on time series. Initially, an ARMA model with periodic variance was proposed by Cao and Wei (1998). Campbell and Diebold (2000) show that conditional heteroskedasticity needs to be added in the model to eliminate the misspecifications observed for US cities. However, the periodic ARMA model was found to be well suited for French cities (see Roustant, 2002) and we will further rely on this approach.

Given a payoff contract based on temperatures, the use of Black and Scholes replication technology is questionable due to the lack of underlying hedging asset[2]. Since for these new products, the market is illiquid and not very transparent, market participants rather use a

[1] However, this time horizon is too long for relying on meteorological forecasts.

[2] One may think to use weather Futures contracts to hedge weather options. However, up to now the weather market is far from being liquid. Unlike electricity there is no spot market for weather quantities.

"marked-to-model" than a "marked-to-market" approach. There are currently a number of pricing methodologies for such contracts, see Carr, Geman and Madan (1999), Davis (2001), Musiela and Zariphopoulou (2001), Schweizer (2001) or Barrieu and El Karoui (2002). In the following, we will rely on a pricing rule, known by actuaries as the "standard deviation principle" (see Goovaerts et al., 1984, or Bühlmann, 1996):

$$P(X) = E\left[\mathrm{X}\right] + \lambda.\sigma\left[\mathrm{X}\right] \qquad (*)$$

where X denotes the expected payoff and $\lambda.\sigma\left[\mathrm{X}\right]$ is known as the safety loading and corresponds to a risk premium. This approach has been used for long by practitioners and has also been thoroughly studied: Moller (2001, 2003a and 2003b) and Schweizer (2001) have considered the pricing of such claims in a financial environment and Hürlimann (2001), Denneberg (1990) compares the standard deviation principle with other pricing approaches. We will thereafter call net premium the first term $E\left[\mathrm{X}\right]$ and for the second one, $\sigma\left[\mathrm{X}\right]$, we will speak of "risk premium".

The mark to model approach results in higher model risk. Reserve policy can be used to cope with model risk. Nevertheless, an important practical point is to look for the sources of model risk: model risk can come through misspecification of the temperature process or from parameter uncertainty[3]. After checking the statistical adequacy of our model (tests, out-of-sample analysis, …), we look for the key parameters with respect to model risk. Thereafter, we use the ARMA temperature model that can be seen as a benchmark. In addition to its simplicity that makes the study feasible, note that it takes into account the major characteristics of temperature (trend, seasonality, seasonality of dispersion, complexity of the dynamics). With this modelling, we assess independently net premium and risk premium uncertainty of weather Futures and options. Let Θ be the vector of parameters involved in the distribution $Q(\Theta)$ of the temperature process. Given some valuation rule such as the standard deviation, the true price is:

$$P = P(\Theta) = E^{Q(\Theta)}\left[\text{discounted payoff}\right] + \lambda.\sigma^{Q(\Theta)}\left[\text{discounted payoff}\right]$$

where $E^{Q(\Theta)}$ and $\sigma^{Q(\Theta)}$ denote the expectation and the standard distribution under $Q(\Theta)$. Since one can only deal with an estimator $\hat{\Theta}$ of Θ, this results in some price uncertainty:

$$\hat{P} = P(\hat{\Theta}) = E^{Q(\hat{\Theta})}\left[\text{discounted payoff}\right] + \lambda.\sigma^{Q(\hat{\Theta})}\left[\text{discounted payoff}\right]$$

Therefore, the price of some weather derivative can itself be seen as a *random variable*[4]. In the following, we will consider confidence intervals of the price estimator as a way to assess this parameter uncertainty. This approach is common in other fields of finance (see Bawa, Brown, Klein, 1979 or Jorion, 1985, for some work related to portfolio management, or Campbell, Lo, MacKinlay, 1997, § 9.3.3, for option pricing). This methodology deals only with parameter uncertainty and assumes that the statistical model is well specified (for a relaxation of this assumption, see Cairns, 2000).

Compared with sensitivity analysis, this statistical approach to model risk assessment takes into account the precision estimation which may vary among parameters. For instance, one may have a good precision in estimating variance and a poor precision with respect to mean parameters.

[3] Since there are no liquid prices we cannot calibrate the parameters from traded prices.

[4] This is not surprising since for illiquid markets, one can only think of a price distribution.

III. PRESENTATION OF DATA AND CONTRACTS.

1. Common specifications

Most of weather contracts are based on temperature. On the CME exchange, the temperature risk for a monthly period is measured by HDD and CDD indexes:

$$HDD = \sum_{t \text{ in month}} (X_{ref} - X_t)^+ \qquad CDD = \sum_{t \text{ in month}} (X_t - X_{ref})^+$$

where X_t is the average temperature at date *t*, defined as the mean of the minimum and maximum temperature observed during the day, and X_{ref} is a baseline temperature (its standard value is 18°C). On the LIFFE-Euronext exchange, temperature indexes are monthly means (or seasonally means) of daily average temperatures. The payoff of a given weather derivative contract is computed from these indexes. This is the main difference with an insurance product: no demonstration of loss is required, the decision of reimbursement and the size of the possible compensations are based on measurement of temperature. It appears that the OTC markets also use such indices. Common specifications of weather derivatives involve the location, the temperature index and the risk exposure period. Now, we explicit our choices for the present study.

- **Location**

As mentioned in the introduction, we focus on the Paris case. For that location indeed, the Gaussian ARMA process used for the temperature modelling (see section IV for details) is rather satisfactory (see Roustant, 2002): unlike the US case (see Campbell and Diebold, 2000), the heteroskedasticity phenomena are not too important for French locations.

- **Temperature index**

We have chosen to work with the Heating Degree-Day (HDD) and Cooling Degree-Day (CDD) indexes:

$$HDD = \sum_{t=u}^{v} (X_{ref} - X_t)^+ \qquad CDD = \sum_{t=u}^{v} (X_t - X_{ref})^+$$

$[u;v]$ is the risk exposure period, X_t denotes the daily average temperature, equal to the mean of the maximum and minimum daily temperatures, and X_{ref} is the standard baseline temperature equal to 18°C. This value is a benchmark in the industry since it is usually considered that heating starts when temperatures go below 18°C.

We could have considered other indexes, such as monthly average of daily temperatures, which are used in the LIFFE-Euronext weather exchange. Actually for months of interest, the two kinds of indexes are closely related, as wee see now. Thus, the methodology presented here applies and we obtain identical conclusions.

- **Risk exposure period**

Usually HDD indexes are computed during the "winter season" from October to April (or November to March, it depends) whereas CDD indexes are used during the "summer season" from May to September. The risk exposure period is mostly considered as the entire season or a specified month. In this paper, we only consider monthly indexes and restrict to the geographic seasons: December to February for winter, June to August for summer. We focus on these contracts for two reasons. First, the months considered are the coldest and the hottest ones, and thus correspond to the most important needs for hedging. In addition, for these months, the results will apply either for CME or LIFFE-Euronext contracts: we argue that for each month in the (geographic) winter season (resp. summer season), we have:

$$HDD \approx \sum_{t=u}^{v} (X_{ref} - X_t) \qquad (1)$$

(resp. $CDD \approx \sum_{t=u}^{v} (X_t - X_{ref})$) ; therefore HDD or CDD indexes are an affine function of the LIFFE-Euronext monthly average of daily temperatures index. This approximation is quite natural since in winter almost all temperatures lie below $X_{ref} = 18°C$. Moreover, assuming that the temperature model of section IV is correct, we can convince ourselves that (1) may hold by simulating a lot of temperature paths and representing the estimated quantiles of the index versus those of its approximation. Such a QQ-plot obtained is drawn in Figure 1. It corresponds to the December case, and was obtained with $R = 10^6$ temperature paths. The perfect observed alignment still remains when we perturb the temperature model's parameters.

In conclusion, we will consider in the present paper monthly risk exposure periods among December, January or February (and thus, relative to HDD index) and June, July and August (relative to CDD index).

2. Weather Futures and options

Temperature-based weather Futures (also called swaps) and options are written on the temperature index, noted TI.

- **Futures pricing**

The Future price is calculated so that the initial value is 0. With our methodology, we can consider that the payoff is the temperature index itself since:

$$0 = E[TI - F] + \lambda.\sigma[TI - F]$$

$$\Leftrightarrow \qquad F = E[TI] + \lambda.\sigma[TI]$$

(F denotes the Future price). From now on, we will adopt this point of view.

- **Weather options**

A fundamental characteristic of an option is its strike price. Firstly, we suggest considering the historical average temperature index's, denoted by K_0. It will be interesting for statistical reasons (see section III, §2.). Of course there is no justification for focusing solely on this particular choice, and finally, we consider five different strikes,

$$K_0 - 60,\ K_0 - 30,\ K_0,\ K_0 + 30,\ K_0 + 60$$

such as two consecutive strikes are separated by a value corresponding to an elevation of one degree per a day.

3. The data

We use daily average temperature from the 1st of January 1979 to the 31st of December 1999, observed at Paris-Montsouris. The source of data is Météo France. In order to facilitate the treatment, we did not consider the 29th of January, which corresponds to remove 5 values per station for a total of 7665 temperatures.

4. Notations

We recall some major notations of the paper:

TI	temperature index
X_t	daily average temperature
NP	net premium: $E[\text{discounted payoff}]$.
RP	"risk premium": $\sigma[\text{discounted payoff}]$.
K_0	historical average temperature index's

IV. PRICE UNCERTAINTY WITH TEMPERATURE MODELLING

As mentioned in section II, we have chosen a simple ARMA model, which captures most of the features of temperature. We study the corresponding price uncertainty and assess marginal influence of each parameter, or group of parameter such as trend, seasonality, volatility or stochastic part. It shows the defects of the temperature model towards weather derivatives pricing issue.

1. The temperature model

The model of temperature takes into account the major characteristics of temperature: seasonality of the values and of the dispersion, quick reversion to the mean, correlations from the days before and today... It has presented by Cao and Wei (1998) or Roustant (2002). It is a linear model with a periodic variance:

$$X_t = m_t + s_t + \sigma_t.Z_t$$

where:

- m_t represents the trend;
- s_t is the seasonal component;
- σ_t is a deterministic and periodic process with an annual periodicity representing the standard deviation of X_t; we assume that $\sigma_t > 0$.
- Z_t is an ARMA process with variance 1 :

$$Z_t = \phi_1.Z_{t-1} + ... + \phi_p.Z_{t-p} + \varepsilon_t + \theta_1.\varepsilon_{t-1} + ... + \theta_q.\varepsilon_{t-q}$$

where (ε_t) is a Gaussian white noise.

Moreover, we assume:

- $m_t = d.t + e$
- $s_t = \sum_{i=1}^{N_f} (a_i.\cos(i.\omega.t) + b_i.\sin(i.\omega.t))$
- $\sigma_t = a + b.\cos(\omega.t) + c.\sin(\omega.t)$

with $\omega = 2\pi/365$.

Choosing such a parametric form allows easy computation of maximum likelihood estimator. For the seasonal component, the choice of frequencies is achieved by means of a preliminary spectral analysis of the normal temperature of each series. In the case of Paris, the discrete curve of the normal temperature is asymmetric which forces the use of at least two frequencies. Finally, we retained the form

$$s_t = \sum_{i=1}^{2} a_i.\cos(i.\omega.t) + b_i.\sin(i.\omega.t)\,.$$

The selection of p and q, the orders of the ARMA model is accomplished by standard procedures (see Brockwell and Davis, 1991), after a preliminary estimation of m_t, s_t and σ_t. It leads to the choice $p = 3$, $q = 0$.

2. Price uncertainty

From now on, we assume that the model of temperature is well specified. Thus, following the approach described above, the price of a weather derivative contract can be computed analytically and is simply a function P of the parameters involved in the model of temperature. If Θ denotes the vector of parameters, and $\hat{\Theta}$ its maximum likelihood estimator, the price $P(\hat{\Theta})$ is the maximum likelihood estimator of $P(\Theta)$. The goal of this paragraph is to quantify the precision of this estimator. This is done in two steps: firstly, an asymptotic pivotal statistic for Θ is explicited; secondly, we derive an approximate distribution for $P(\hat{\Theta})$. Price uncertainty is assessed as an interquantile interval. Moreover, we argue that, when the payoff is approximately a linear function of temperature, the result can be fully interpreted as a confidence interval. This is the case of the net premium of Futures corresponding to "non- intermediate" months considered here, such as winter or summer months.

2.1. Statistical properties of the temperature model's parameters estimator

The model used for temperature is roughly an ARMA model with a deterministic pattern including trend and seasonality, and a deterministic correction for the seasonal volatility. We recall that for an ARMA process, we have the following result. If Θ is the parameters vector and $\hat{\Theta}$ the maximum likelihood estimator of Θ, then $\hat{\Theta}$ is asymptotically Gaussian (see Gouriéroux and Montfort, 1997): $\sqrt{n}\left(\hat{\Theta} - \Theta\right) \xrightarrow[n \to +\infty]{} N(0; \Gamma(\Theta))$.

Here, we assume that this results still holds for the periodic ARMA process. This is not unreasonable for two reasons. Firstly, the additional parametric forms m_t, s_t and σ_t for trend, seasonality and volatility are all deterministic. Secondly, remark that this asymptotical result may give a good proxy for the exact distribution of $\sqrt{n}\left(\hat{\Theta} - \Theta\right)$ since the sample size is rather large ($n = 7.665$); Now, we can check *a posteriori* the normality property for the distribution of $\hat{\Theta}$. Indeed, by using a simulation technique, as shown in the appendix, one can obtain independent realisations of $\hat{\Theta}$. Then, standard normality tests can be used. We have studied both the normality of the marginal distributions using the Kolmogorov test and the normality of the whole distribution, with the skewness and kurtosis tests, as presented in (Lütkepohl, 1993, § 4.5., formulas 4.5.4., 4.5.5. and 4.5.8). The results (see Table 2) show that the distributional assumption for $\hat{\Theta}$ is satisfactory[5].

[5] In particular, the difficulty for the trend parameters is not general: results obtained with an American temperature series show it. Rather, it may indicate that the asymptotic normality is not totally reached yet.

The same simulated realisations can be used to estimate the covariance matrix $\Gamma(\Theta)$. Thus, the normality of the temperature model's parameter estimator can be used to derive confident intervals for the weather derivatives products[6]. This is the aim of next paragraph.

Moreover, let us mention that the expression of the covariance matrix provides us with information about the temperature model itself. For easier legibility, we have reported details and results in the appendix (see Table 1 for the covariance matrix). The most striking one is the significance of the growing trend. Indeed, even if the estimation results of literature have exhibited a growing trend (explained by the global warming and urbanisation phenomena), this one is rather small and might have been non significant at all, because of estimation errors. We can see that it is not the case.

2.2. Assessment of price uncertainty.

We discuss two methodologies for the assessment of Future price uncertainty the first one based on asymptotic expansions and the second one based on simulation.

- **Asymptotic methodology.**

From the assumption above, we can easily derive an asymptotic pivotal statistic of $P(\Theta)$, by using a first-order Taylor expansion (see Campbell, Lo and MacKinlay, 1997, section A.4. of the appendix). We then have:

$$\sqrt{n}\left(P(\hat{\Theta})-P(\Theta)\right)\xrightarrow[n\to+\infty]{} N(0;v_P(\Theta)) \tag{1}$$

where $v_P(\Theta)=\frac{\partial P}{\partial\Theta'}\Gamma(\Theta)\frac{\partial P}{\partial\Theta}$.

If we can compute the expression above, then we can estimate $v_P(\Theta)$ in a natural way by replacing Θ by its sample estimation, and deduce an asymptotic confidence interval for $P(\Theta)$.

The major drawback with this methodology is that we have no idea of the accuracy of the asymptotic distribution of $P(\hat{\Theta})$. Thus, it is useful to look for the shape of the exact distribution and to precise the situations where the two ones merge.

- **Simulation methodology.**

We get an approximation of the exact distribution of $P(\hat{\Theta})$ by using a simulation technique:

– simulate independently $\hat{\Theta}^1$, ..., $\hat{\Theta}^R$ with the Gaussian distribution of covariance matrix $\Gamma(\Theta)$ obtained previously;

– calculate $P\left(\hat{\Theta}^1\right)$, ..., $P\left(\hat{\Theta}^R\right)$.

Therefore price uncertainty is assessed by two quantiles of $P(\hat{\Theta})$ estimated with the simulated sample $P\left(\hat{\Theta}^1\right)$, ..., $P\left(\hat{\Theta}^R\right)$.

- **Discussion**

Both methodologies lead to approximate confidence intervals. However, the uncertainty may be assessed in a more realistic manner if the exact distribution departs from the asymptotic Gaussian one. In that case, we will prefer following the simulation methodology.

---

[6] As we will see, the asymptotic part of the result is not strictly required.

- **Linear payoff.**

In last paragraph, we have noticed the importance of situations in which payment is a Gaussian random variable. In practise, this happens when the payoff is a linear function; in our purpose, this is the case for the net premium of a Future contract for any "non-intermediate" months, such as winter or summer months. Indeed, for the months considered here, the temperature index is an affine combination of the model's parameters. By linearity of the payoff, $P(\hat{\Theta})$ then reduces to an affine combination of the model's parameters. Since the parameters estimator is approximately normally distributed, we conclude that $P(\hat{\Theta})$ is approximately normally distributed too.

- **Results**

Some results are indicated in Table 3, Table 4 and Table 5. The conclusions are quite different whether they concern Futures or options contracts. For Futures, price uncertainty is reasonable and may even be acceptable in some cases, as for a winter month HDD contract with an order of magnitude of 5%. On the other hand, we observe large price uncertainty in the options case: the net premium uncertainty is about 50% when the strike is the historical average temperature index's and can increase dramatically for higher strike prices (actually, the quality of estimation seems to decrease with the estimated price).

To go further in the understanding of these results, we now focus on local influence on prices of parameters, or groups of parameters.

2.3. Assessment of local price uncertainty

Up to the present time, we have assessed price uncertainty due to estimation error of all parameters. But can we quantify the impact of estimation error around trend only (for instance) on the weather derivative price? We will speak of *local* uncertainty as opposed to (global) price uncertainty studied before. Denote by Θ_1 the set of parameters we want to assess influence (in our example, Θ_1 consists of the two parameters modelling trend) and Θ_2 the other ones; let $\hat{\Theta}_1$ and $\hat{\Theta}_2$ the corresponding (maximum likelihood) estimators. Then the issue can be addressed by estimating the conditional distribution:

$$P\left(\hat{\Theta}_1\right)\Big|\hat{\Theta}_2 = \Theta_2$$

Therefore local price uncertainty can be assessed exactly in the same way as for price uncertainty. Indeed, assume that

$$\sqrt{n}\left(\hat{\Theta}-\Theta\right)\xrightarrow[n\to+\infty]{} N(0;\Gamma(\Theta))$$

Now introduce the partition in $\Gamma(\Theta)$ corresponding to $\Theta=\left(\Theta_1;\Theta_2\right)$: $\Gamma=\begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix}$; hence Γ_{ij} represents the (asymptotic) matrix of the covariances between $\sqrt{n}\left(\hat{\Theta}_i-\Theta_i\right)$ and $\sqrt{n}\left(\hat{\Theta}_j-\Theta_j\right)$. Then we have (see Saporta, § 4.2.4):

$$\sqrt{n}\left(\hat{\Theta}_1-\Theta_1\Big|\hat{\Theta}_2=\Theta_2\right)\xrightarrow[n\to+\infty]{} N(0;\Gamma_{11}-\Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21})$$

Therefore, using again a first Taylor expansion (see § 2.2. *Asymptotic methodology*), we get:

$$\sqrt{n}\left(P(\hat{\Theta}_1)-P(\Theta_1)\Big|\hat{\Theta}_2=\Theta_2\right)\xrightarrow[n\to+\infty]{} N(0;v_{P|\Theta_2}(\Theta_1))$$

with $v_{P|\Theta_2}(\Theta_1) = \frac{\partial P}{\partial \Theta_1{}'}\left(\Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}\right)\frac{\partial P}{\partial \Theta_1}$.

In practise, as discussed before, we do not use last result but prefer resampling in the conditional asymptotic distribution of $\hat{\Theta}_1$ to get an approximation of the exact distribution of $P(\hat{\Theta}_1)\Big|\hat{\Theta}_2 = \Theta_2$.

- **Results**

Figure 2 and Figure 4 show the marginal influence of groups of parameters. They represent simulated samples (obtained with $R = 10^4$ simulations) of the (entire) price distribution and the price conditional distributions corresponding to trend, seasonality, volatility or autoregressive part. Before discussing the results, remark that in the Future case, only two groups of parameters come into account in the net premium and the risk premium expression (see Appendix A.1.) - trend and seasonality only for the former, dispersion and stochastic part only for the latter. Now, we clearly observe that the largest uncertainties on prices are essentially caused by trend and seasonality estimation errors. We can even say that *both* trend and seasonality are responsible for price uncertainty since the involved parameters are nearly independent (see the correlation matrix, Table 1).

In order to assess the local influence of each parameter, we compare prices conditional distributions corresponding to a sole parameter (Figure 3 and Figure 5). We conclude that trend slope is the most harmful parameter. To obtain this result, a change of parameter has been necessary; indeed in the basic form the two trend parameters are highly negatively correlated and it would have been impossible to dissociate the specific influence of only one of them. On the other hand, with the new parameterisation, trend slope is nearly independent of all other parameters (see Appendix A.3.). This confirms the idea that the trend modelling by a straight line is too rough and, in particular, that a small deviation of the slope should cause a large deviation on the temperature level and then on the (model) prices. However the lack of accuracy is not due solely to this parameter as we have seen on Figure 3 and Figure 5.

Finally, we see that the temperature model imperfections that are most harmful for weather derivatives prices, come from the modelling of the deterministic parts relative to the process mean, that is: trend and seasonality.

V. CONCLUSION

Motivated by the development of statistical modelling for weather derivatives, we have addressed the issue of price uncertainty coming from risk estimation in a given model. We have assessed price uncertainty obtained with a Gaussian ARMA process (with periodic volatility). We observe dramatic uncertainty around options prices. After searching for the source of that uncertainty, we conclude that the parameters of the process mean - that is: trend and seasonality parameters, are essentially responsible for it.

These conclusions indicate that some efforts should be done on the modelling of trend and seasonality. For instance, new models could allow stochastic trend and seasonality. Of course the improvements of the stochastic part, such as the incorporation of conditional heteroskedasticity in the model (Campbell and Diebold, 2000) or long memory effects (Brody, Syroka and Zervos, 2001), should also be useful, especially for options.

APPENDIX

A.1. Analytical expressions of options net premiums and risk premiums

By an immediate computation, one can show the following:

Lemma. *Let K be a real, and X a Gaussian random variable,* $X \sim N\left(\mu,\sigma^2\right)$. *Then, with* $Y=\left(K-X\right)^+$ *(resp.* $Y=\left(X-K\right)^+$*), we have:*

$$E\left[Y^+\right]=\sigma\left(L.N(L)+\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{L^2}{2}\right)\right)$$

$$E\left[\left(Y^+\right)^2\right]=\sigma^2\left(\left(1+L^2\right).N(L)+\frac{L}{\sqrt{2\pi}}\exp\left(-\frac{L^2}{2}\right)\right)$$

where $L=\frac{K-\mu}{\sigma}$ *(resp.* $L=-\frac{K-\mu}{\sigma}$*.), and N(.) denotes the standard normal cumulative function.*

Application. For instance consider the case of HDD options. Denote by [u; v] the period of risk exposure, and $\left(X_t\right)$ the temperature process. Then, for the months considered here, the HDD index equals approximately

$$HDD \approx \sum_{t=u}^{v}(X_{ref}-X_t)$$

where X_{ref} is the standard threshold, that is $X_{ref}=18°C$. Since $\left(X_t\right)$ is a Gaussian process, we deduce that

$$HDD \sim N\left(l.X_{ref}-\sum_{t=u}^{v}(m_t+s_t);\rho'.\Gamma_Z.\rho\right)$$

with notations of section III and where, in addition, $l=v-u+1$ (the period length), ρ is the vector $l\times 1$, $\rho=(\rho_u;\rho_{u+1};...;\rho_v)$ and Γ_Z the $l\times l$ autocorrelations matrix of the autoregressive process Z, $\left(\Gamma_Z\right)_{i,j}=\text{cov}\left(Z_0;Z_{|i-j|}\right)$, $1\le i,j\le l$. Then, the lemma applies.

A.2. Statistical properties of the temperature model parameters

We consider the temperature model

$$X_t=m_t+s_t+\sigma_t.Z_t,$$

where $Z_t=\phi_1.Z_{t-1}+...+\phi_p.Z_{t-p}+\varepsilon_t$, and the residuals process $\left(\varepsilon_t\right)$ is a Gaussian white noise. Denote by Θ the parameter vector, and $\hat{\Theta}$ the maximum likelihood estimator. In §IV.2, we have assumed that $\sqrt{n}(\hat{\Theta}-\Theta)\xrightarrow[n\to+\infty]{}N\left(0,\Gamma(\Theta)\right)$. Now, we show below how to obtain independent realisations Θ^{*1},..., Θ^{*R} of $\hat{\Theta}$ without using the preceding assumption, by a simulation technique. These ones are used to check the distributional assumption of $\hat{\Theta}$ (see Table 2), and lead to the standard estimation of $\Gamma(\Theta)$ by the empirical covariance matrix computed from the Θ^{*r} (see results below):

$$\frac{\hat{\Gamma}_{i,j}}{n} = \frac{1}{R}\sum_{r=1}^{R}(\Theta_i^{*r} - \overline{\Theta}_i)(\Theta_j^{*r} - \overline{\Theta}_j)$$

with $\overline{\Theta}_i = \frac{1}{R}\sum_{r=1}^{R}\Theta_i^{*r}$ (Θ_i^{*r} denotes the i-th coordinate of Θ^{*r}).

Independent realisations from the distribution of $\hat{\Theta}$ can be obtained from the data by simulation in the standard way, as shown in (Efron and Tibshirani, 1986, §6). Practically, for $r = 1,...,R$, we do the following operations:

– simulate independently $\varepsilon_1^*,...,\varepsilon_n^*$ from the estimated Gaussian distribution of residuals.

– reconstitute the corresponding temperature path $x_1^*,...,x_n^*$ using the temperature model with the *initial* parameter estimation :

$$x_t^* = m_t + s_t + \sigma_t.z_t^*,$$

with $z_t^* = \hat{\phi}_1.z_{t-1}^* + ... + \hat{\phi}_p.z_{t-p}^* + \varepsilon_t^*$ and $z_0^* = z_{-1}^* = ... = z_{1-p}^* = 0$;

– calculate the maximum likelihood estimate Θ^{*r} from the "new data" $x_1^*,...,x_n^*$.

Actually, we make a statistical error by replacing the true distribution of residuals by its estimated one, and the true parameter vector by its (maximum likelihood) estimation. Nevertheless, it may be small because of the large sample size, n = 7665. Thus we can consider that Θ^{*1},..., Θ^{*R} are independent realisations of $\hat{\Theta}$.

- **Numerical result**

Below, we give the standard deviation of estimation of the parameters, and the correlation matrix computed from the covariance matrix obtained using the aforementioned simulation technique with $R = 2.000$ simulations. To improve the quality of estimation, a new parameterisation was adopted for trend. Parameters have been gathered according to their belonging to the characteristics of the model: trend, seasonality, dispersion and autoregressive part. Half of the matrix is represented for legibility reasons; one will deduce the other side by symmetry.

Here are some remarks that can be made.

➢ About the estimation of the trend, we see that the estimated slope is positive in a significant manner, though the estimation error is large; thus the trend is significantly increasing. Moreover the estimators of the slope *d* and the constant *e* are highly negatively correlated (not given here). Indeed, overestimating the slope result in underestimating the constant, as increasing the inclination of a mobile bar fixed at its middle result in lowering the lowest end. One can reduce dramatically this correlation by replacing the constant *e* by $e' = d.\bar{t} + e$ where $\bar{t}$ is the mean time. This is what we have done.

➢ Moreover, the estimators of *a* and ϕ_1 are highly positively correlated. We can understand it by approximating *a* by the standard deviation of the process σ_t and approximating Z_t by an AR(1) process, so that $X_t \approx m_t + s_t + a.Z_t$ with $Z_t \approx \phi_1.Z_{t-1} + \varepsilon_t$. Hence we remark that, intuitively, overestimating ϕ_1 results in decreasing the dispersion of Z_t which leads to overestimate *a* to "fit" the dispersion of the data.

Eventually, the estimation of the coefficients of the autoregressive process and their correlations are very close to what we obtain by estimating the coefficients of an AR(3) model using the de-trended and deseasonalised data (but with a constant dispersion). The maximum likelihood estimation procedure of such a model is a simple least squares estimation procedure, whereas the estimation of the temperature model, which is also a maximum likelihood procedure, is nearly equivalent to use least squares weighted by σ_t. The fact that the weighting is not too significant may explain the closeness of the results. The introduction of a periodic variance is especially useful to obtain closer to reality errors (Roustant, 2002).

Table 1. Parameters of the temperature model: estimated values, standard deviations and correlation matrix

Paris-Montsouris

			(Trend)		*(Seasonality)*				*(Dispersion)*			*(AR process)*		
θ	$\hat{\theta}_{emv}$	σ	*d*	*e'*	a_1	b_1	a_2	b_2	*a*	*b*	*c*	ϕ_1	ϕ_2	ϕ_3
d	1.85*	.48*	1											
e'	12.1	.11	.00	1										
a_1	-7.39	.15	.01	.07	1									
b_1	-2.67	.15	.06	.04	-.01	1								
a_2	-0.14	.15	-.01	-.01	.03	-.05	1							
b_2	0.74	.15	.03	.02	.04	.01	.00	1						
a	3.35	.05	-.00	.01	.01	.00	.02	.03	1					
b	0.11	.04	-.01	.00	.03	.01	-.02	.01	.02	1				
c	0.15	.04	-.02	.02	-.02	-.02	.02	.00	.10	-.03	1			
ϕ_1	0.92	.01	-.01	.02	-.01	.00	.01	.02	.57	-.01	.03	1		
ϕ_2	-0.19	.01	.01	.01	.03	.01	-.00	-.01	-.08	.01	.03	-.68	1	
ϕ_3	0.06	.01	.02	.03	-.02	-.01	.00	.01	.06	.02	-.01	.16	-.68	1

* : $\times 10^{-4}$

Table 2. Results of the normality tests for the distribution of $\hat{\Theta}$

	Paris	Chicago
Marginal distributions	Kolmogorov statistic	
$\hat{d}$	* 0.0248	0.0160
$\hat{e}$	* 0.0228	0.0167
$\hat{a}_1$	0.0101	0.0152
$\hat{b}_1$	0.0169	0.0129
$\hat{a}_2$	0.0171	
$\hat{b}_2$	0.0137	
$\hat{a}$	0.0165	0.0155
$\hat{b}$	0.0105	0.0146
$\hat{c}$	0.0095	0.0094
$\hat{\phi}_1$	0.0158	0.0193
$\hat{\phi}_2$	0.0133	0.0135
$\hat{\phi}_3$	0.0141	0.0162
Multidimensional distribution	p-value	
Skewness test	0.5955	0.6258
Kurtosis test	0.0517	0.0879
Joint test	0.1497	0.2227

This table gives the results of normality tests for samples of length $R = 2.000$, obtained by simulation on temperature data relative to Paris-Montsouris and O'Hare Airport (Chicago) [for the latter, data are available freely on the Chicago Mercantile Exchange website]. For Kolmogorov test, with 5% confidence level, normality is rejected here when the statistic value is superior to 0.0200; corresponding cases are indicated by a star.

REFERENCES

Barrieu P. and El Karoui N. (2002), *Optimal design of derivatives in illiquid markets*, Quantitative Finance, 2, 181-188.

Bawa V.S., Brown S.J. and Klein R.W. (1979), *Estimation Risk and Optimal Portfolio Choice*, Studies in Bayesian Econometrics Bell Laboratories Series, New York: North Holland.

Brockwell P.J. and Davis R.A. (1991), *Time Series: Theory and Methods*, Springer Verlag.

Brody D.C., Syroka J. and Zervos M. (2001), *Dynamical Pricing of Weather Derivatives*, working paper.

Bühlmann H. (1996), *Mathematical Methods in Risk Theory*, Springer-Verlag, second printing.

Cairns A.J.G. (2000), *A discussion of parameter and model uncertainty in insurance*, Insurance: Mathematics and Economics, 27, 313-330.

Campbell J.Y., Lo A.W. and MacKinlay A.C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Campbell S. and Diebold F.X. (2000), *Weather Forecasting for Weather Derivatives*, working paper.

Cao M. and Wei J. (1998), *Pricing Weather Derivatives: an Equilibrium Approach*, working paper.

Carr P., Geman H. and Madan D. (1999), *Pricing and Hedging in Incomplete Markets*, working paper.

Davis M. (2001), *Pricing weather derivatives by marginal value*, Quantitative Finance, 1, 1-4.

Davinson A.C. and Hinkley D.V. (1997), *Bootstrap Methods and their Applications*, Cambridge University Press.

Denneberg D. (1990), *Premium calculation: why standard deviation should be replaced by absolute deviation*, Astin Bulletin, 20, 181-190.

Dornier F. and Queruel M. (2000), *Caution to the Wind Risk,* Energy and Power Risk Management.

Efron B. and Tibshirani R. (1986), *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Accuracy*, Journal of Statistical Science, 1(1), 54-77.

Goovaerts M.J., DeVylder F. and Haezendonck J. (1984), *Insurance Premiums*, Elsevier.

Gouriéroux C. and Montfort A. (1997), *Time Series and Dynamics Models*, Cambridge University Press.

Hürlimann W. (2001), *Distribution free comparison of pricing principles*, Insurance: Mathematics and Economics, 28, 351-360.

Jorion P. (1985), *International Portfolio Diversification with Estimation Risk*, Journal of Business, 58(3), 259-278.

Lütkepohl H. (1993), *Introduction to Multiple Time Series Analysis*, Springer Verlag.

Moller T. (2001), *On transformations of actuarial valuation principles*, Insurance: Mathematics and Economics, 28, 281-303.

Moller T. (2003), *Indifference pricing of insurance contracts in a product space model: applications*, Insurance: Mathematics and Economics, 32, 295-315.

Moller T. (2003), *Indifference pricing of insurance contracts in a product space model*, Finance and Stochastics, 7, 197-217.

Moréno M. (2000), *Riding the Temp*, Weather Derivatives, FOW Special Supplement.

Musiela M. and Zariphopoulou T. (2001), *Pricing and Risk Management of Derivatives Written on Non-Traded Assets*, working paper.

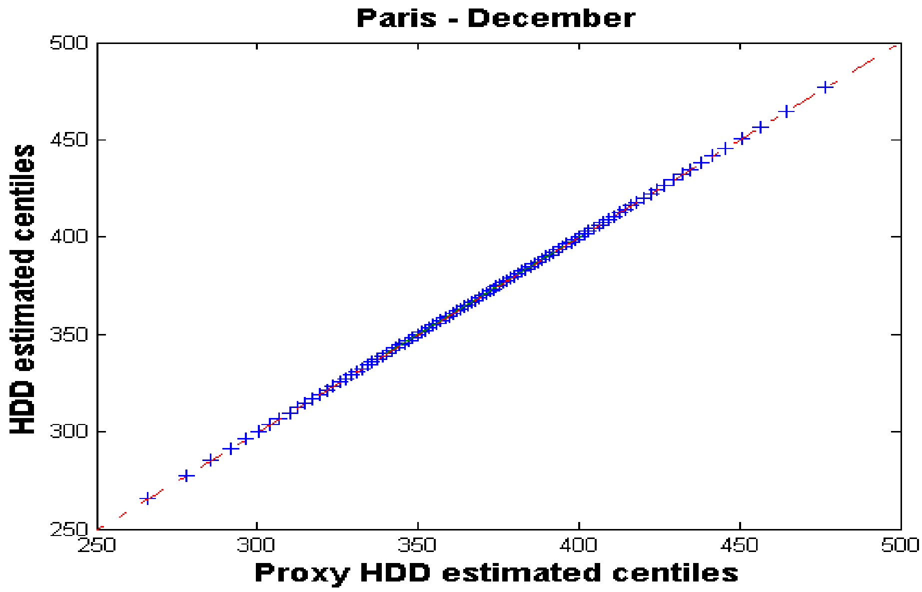
Roustant O. (2002), *Une application de deux modèles économétriques de température à la gestion de risques climatiques (first part)*, Banque & Marchés, 58, 22-29.

Roustant O. (2002), *Une application de deux modèles économétriques de température à la gestion de risques climatiques (second part)*, Banque & Marchés, 59, 36-44.

Saporta G. (1990), *Probabilités, Analyse des données et Statistiques*, Technip.

Schweizer M. (2001), *From actuarial to financial valuation principles*, Insurance : Mathematics and Economics, 28, 31-47.

Figure 1. Validity of the linear approximation for the temperature indexes.



QQ-plot of the estimated centiles of the December HDD index (at Paris) versus those of its linear approximation, obtained with 10^6 simulations.

Table 3. Future price uncertainty.

December HDD Future. Paris.

	Net premium	Risk premium
Estimated value	370.0	49.1
Uncertainty	[351.6; 388.9]	[45.5; 52.2]
Relative uncertainty* (%)	[-5.2; 5.0]	[-7.3; 6.3]

July CDD Future. Paris.

	Net premium	Risk premium
Estimated value	88.3	46.1
Uncertainty	[70.1; 106.2]	[43.0; 48.9]
Relative uncertainty* (%)	[-20.5; 20.4]	[-6.6; 6.0]

* *around the estimated value.*

Table 4. Option price uncertainty.

Call option on Future December HDD, with strike price K_0. Paris.

	Net premium	Risk premium
Estimated value	14.5	24.9
Uncertainty	[8.1; 23.6]	[18.3; 31.3]
Relative uncertainty* (%)	[-44.2; 62.3]	[-26.2; 25.8]

Call option on Future July CDD, with strike price K_0. Paris.

	Net premium	Risk premium
Estimated value	22.6	29.6
Uncertainty	[13.7; 34.2]	[23.2; 35.3]
Relative uncertainty* (%)	[-39.2; 51.0]	[-21.6; 19.2]

K_0 is the historical average temperature index's.

** around the estimated value.*

Table 5. Option price uncertainty versus strike price.

Net premium of some December HDD call options. Paris.

Strike price	Estimated value	Uncertainty	Relative uncertainty* (%)
$K_0 - 60$	53.0	[38.4; 69.6]	[-27.5; 31.2]
$K_0 - 30$	30.4	[19.5; 44.1]	[-35.9; 44.5]
K_0	14.5	[8.1; 23.6]	[-44.2; 62.3]
$K_0 + 30$	5.5	[2.5; 10.5]	[-54.6; 89.6]
$K_0 + 60$	1.6	[0.6; 3.7]	[-62.6; 126.8]

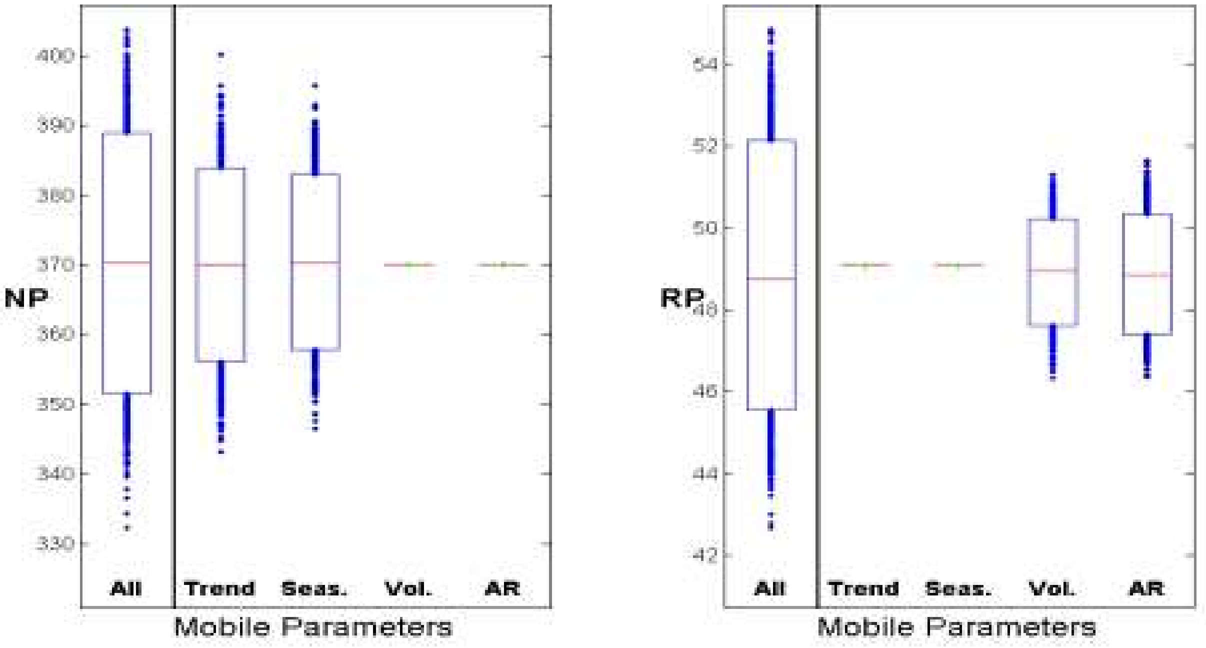
Risk premium of the some December HDD call options. Paris.

Strike price	Estimated value	Uncertainty	Relative uncertainty* (%)
$K_0 - 60$	42.5	[37.6; 46.2]	[-11.5; 8.7]
$K_0 - 30$	34.8	[28.3; 40.3]	[-18.8; 15.5]
K_0	24.9	[18.3; 31.3]	[-26.2; 25.8]
$K_0 + 30$	15.1	[9.7; 21.2]	[-35.2; 40.4]
$K_0 + 60$	7.8	[4.4; 12.2]	[-42.4; 56.7]

K_0 *is the historical average temperature index's. Here,* $K_0 = 381.1$.

* *around the estimated value.*

Figure 2. Future price uncertainty - Local price uncertainty due to the model components.

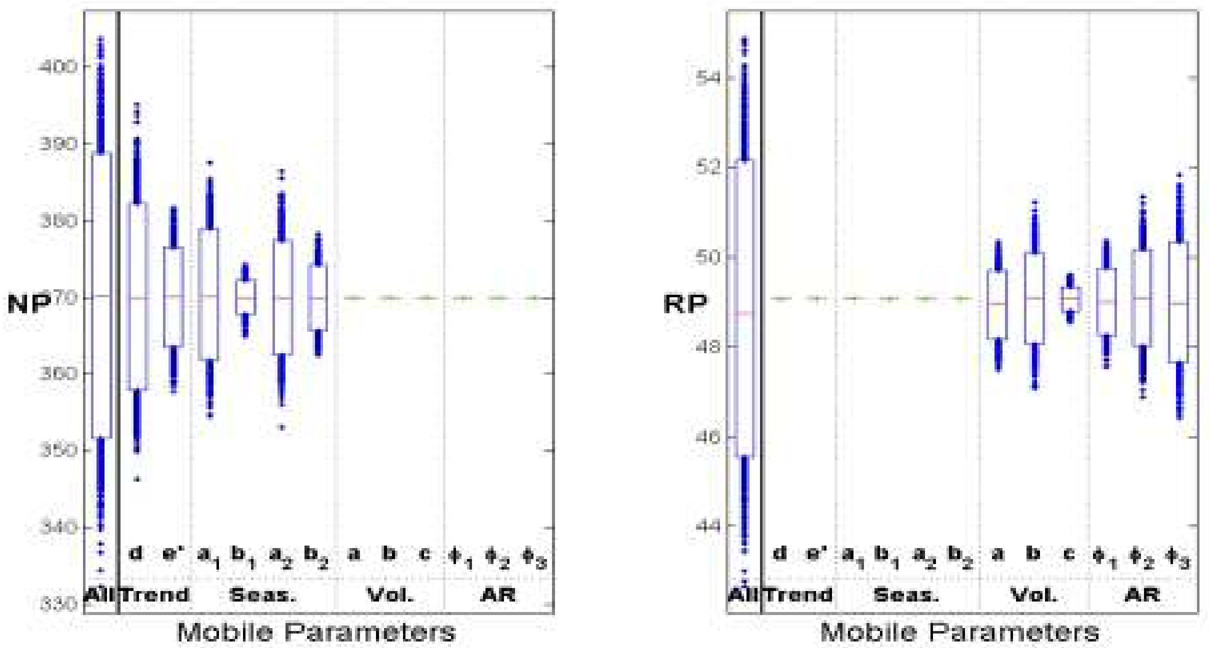


Simulated samples of the price distribution and conditional distributions***
of the December HDD Future, Paris.

* NP: net premium – RP: risk premium.

** In the conditioning, all parameters are fixed except those involved in the specified group of parameters: trend, seasonality, volatility or autoregressive part (AR).

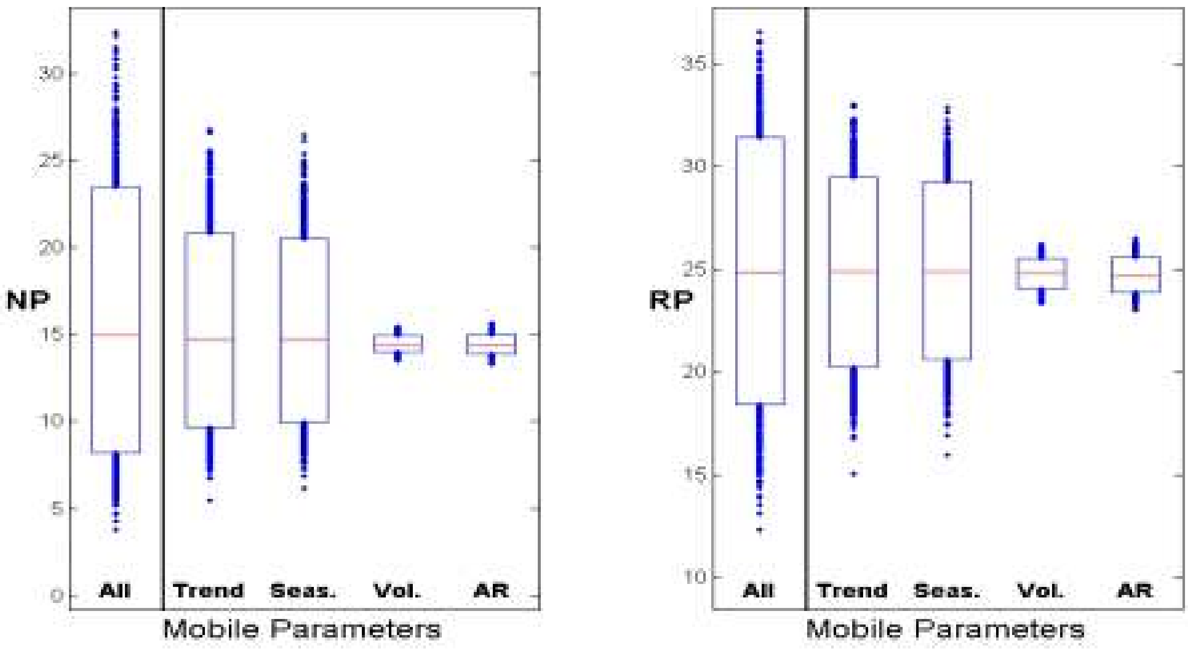
Figure 3. Future price uncertainty - Local price uncertainty due to each parameter.



Simulated samples of the price distribution and conditional distributions***
of the December HDD Future, Paris.

* NP: net premium – RP: risk premium.
** In the conditioning, all parameters are fixed except the one indicated.

Figure 4. Option price uncertainty - Local price uncertainty due to the model components.



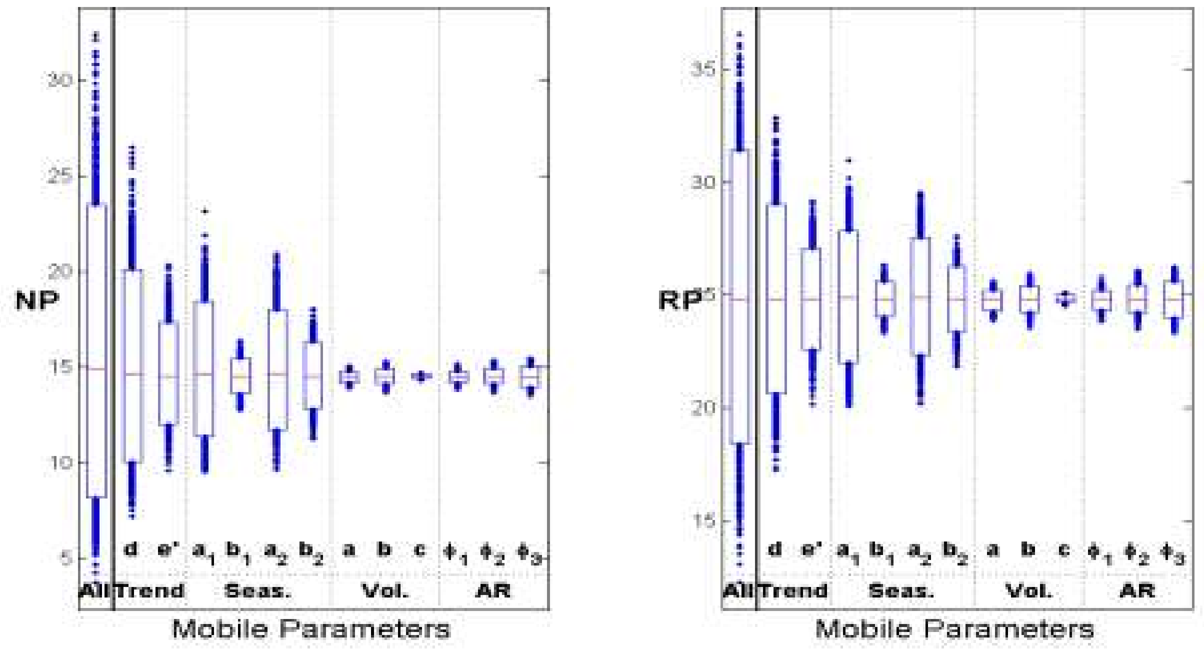
Simulated samples of the price* distribution and conditional distributions** of the call option on the December HDD Future with strike price K_0 , Paris.

K_0 *is the historical average temperature index's.*

** NP: net premium – RP: risk premium.*

*** In the conditioning, all parameters are fixed except those involved in the specified group of parameters: trend, seasonality, volatility or autoregressive part (AR).*

Figure 5. Option price uncertainty - Local price uncertainty due to each parameter.



Simulated samples of the price* distribution and conditional distributions** of the call option on the December HDD Future with strike price K_0, Paris.

K_0 is the historical average temperature index's.

** NP: net premium – RP: risk premium.*

*** In the conditioning, all parameters are fixed except those involved in the specified group of parameters: trend, seasonality, volatility or autoregressive part (AR).*