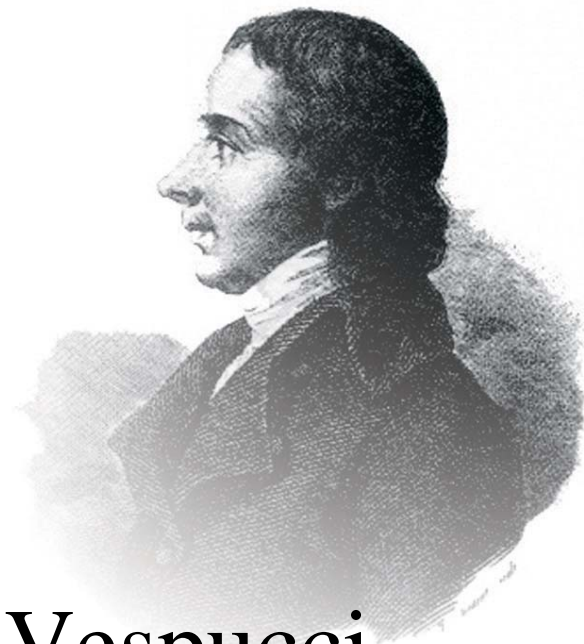


Rapporto n. 4 2006

dmsia  unibg.it



F. Maggioni, M.T. Vespucci,  
E. Allevi, M.I. Bertocchi,  
M. Innorta

A stochastic optimization  
model for gas sale company  
using mean reverting process  
modelling of the temperature

*Serie Ricerca*

Dipartimento  
di Matematica, Statistica,  
Informatica e Applicazioni  
“Lorenzo Mascheroni”

UNIVERSITÀ DEGLI STUDI DI BERGAMO



# A stochastic optimization model for gas sale company using mean reverting process modeling of the temperature

F. Maggioni<sup>2</sup>, M.T. Vespucci<sup>3</sup>, E. Allevi<sup>1</sup>, M.I. Bertocchi<sup>2</sup> and M. Innorta<sup>3</sup>.

**Abstract.** The paper deals with a new stochastic optimization model, named OMoGaS-SV (Optimization Modelling for Gas Seller-Stochastic Version), to assist companies dealing with gas retail commercialization. Stochasticity is due to the dependence of consumptions on temperature uncertainty. Two different models for modelling temperature scenarios are compared. Due to nonlinearities present in the objective function, the model can be classified as an NLP mixed integer model, with the profit function depending on the number of contracts with the final consumers, the typology of such consumers and the cost supported to meet the final demand. Constraints related to a maximum daily gas consumption, to yearly maximum and minimum consumption in order to avoid penalties and to consumption profiles are included. The results obtained by the stochastic version give clear indication of the amount of losses that may appear in the gas seller's budget.

**Key Words.** Gas sale company, tariff components, mean reverting process, stochastic programming.

## 1 Introduction

Starting in 1999 the Italian Natural Gas market has been undergoing a liberalization process aiming at promoting competition and efficiency, while ensuring adequate service quality standards. Timings and methods for the internal gas market liberalization have been introduced following the European Gas Directive; the roles of different segments of the natural gas “chain ” have been identified and defined, such as import, production, export, transportation and dispatching, storage, distribution and sale. In particular, the principle has been introduced of unbundling among supply and transport/storage and among distribution and selling. Before liberalization there was a national monopolistic operator, for all activities related to supply, transport, storage and wholesale commercialization, and local monopolistic operators, for distributing and selling to final consumers. After liberalization the following operators run different activities:

---

<sup>1</sup>Department of Quantitative Methods, Brescia University, Contrada S. Chiara, 50, Brescia 25122, Italy.

<sup>2</sup>Department of Mathematics, Statistic, Computer Science and Applications, Bergamo University, Via dei Caniana 2, Bergamo 24127, Italy.

<sup>3</sup>Department of Engineering and Management, Bergamo University, Via Marconi 5, Dalmine 24044, Italy.

- shippers: production/import, re-gasification and wholesale commercialization;
- national distributor: transport on national network and storage;
- local distributors: transport on local network;
- selling companies: purchase gas from shippers and sell to final consumers.

In 2003 the Italian Regulatory Authority for Electricity and Gas, see [9], defined consumption classes, on the basis of gas consumption in the thermal year, and introduced a new gas tariff, in order to guarantee small consumers' protection by applying the transparency principle in the pricing mechanism. The new tariff is based on a detailed splitting in different components, whose values are periodically revised, and represents a maximum price to be applied to small consumers.

In a previous paper, see Allevi *et al.*, (2006) [2], a deterministic optimization model has been developed to assist companies dealing with retail commercialization. For each citygate, the gas seller has to decide the customer portfolio, i.e. the number of final customers to supply in each consumption class, and the sell prices to apply to each consumption class. Indeed, different customer portfolios determine different citygate consumption patterns, which shippers refer to when setting the gas price to be paid by the gas seller for the citygate supply. For each thermal year and each citygate there is a contract between shipper and gas seller setting:

- the gas volume required by gas seller for next thermal year;
- the gas volume required in particular in winter months;
- the maximum daily consumption (capacity) requested by gas seller;
- the purchase price fixed by the shipper.

In the contract it is also specified how to compute penalties, due by gas seller if daily consumption exceeds daily capacity.

In this paper we introduce stochasticity in the model due to the influence of temperature on consumptions. For domestic customers, using gas either only for cooking or for cooking and heating, and for commercial activities and small industries, gas consumption in winter months strongly depends on the weather conditions: this fact is taken into account in the model, by including a mean reverting process modeling temperature, which gas consumption depends on. This model is presented in section 2. In section 3 the stochastic model, named OMoGaS-SV, is presented and in section 4 numerical results related to a case study are reported and discussed.

## 2 The stochastic temperature model

In this section we introduce a stochastic model describing the temperature variations along the months in a year time. We start with some definitions about temperature:

**Definition 1** Given a weather station, let  $T_\mu^{max}$  and  $T_\mu^{min}$  denote the maximum and the minimum temperatures (in Celsius degrees) measured in day  $\mu$ , respectively. We define the **mean temperature** of day  $\mu$  as

$$T_\mu = \frac{T_\mu^{max} + T_\mu^{min}}{2} . \quad (1)$$

**Definition 2** Let  $T_\mu$  denote the mean temperature of day  $\mu$ . We define **Heating Degree Days** ( $HDD_\mu$ : measure of cold in winter) and **Cooling Degree Days** ( $CDD_\mu$ : measure of heat in summer) respectively as

$$HDD_\mu = \max \{18 - T_\mu, 0\} , \quad (2)$$

$$CDD_\mu = \max \{T_\mu - 18, 0\} . \quad (3)$$

For a given day HDD and CDD are the numbers of degrees of deviation from a reference temperature level in Bergamo ( $18^\circ \text{C}$ ). The name “heating degree days” refers to the fact that if temperature is below  $18^\circ \text{C}$  people tend to use more energy to heat their homes; the name “cooling degree days” refers to the fact that if temperature is above  $18^\circ$  people start turning their air conditioners on. Typically the HDD season is from November to March, whereas the CDD season is from May to September. April and October are often referred to as “shoulder months”.

We have a database of temperatures measured in Bergamo in the last 12 years (1/01/1994-30/11/2005). The database consists of daily minimum and maximum temperatures, from which average daily temperatures are computed using (1). Due to the cyclical nature of the temperature process we find that historical data give a reasonable idea of the temperature level in the future. In Figure 1 we have plotted the daily mean temperatures at Bergamo for the 12 years; from the picture is clear that the temperature process should be a mean reverting process, reverting to some cyclical function.

The histogram of the daily temperature differences in Bergamo (1994 – 2005) is given by Figure 2. Clearly, the daily differences approximate a normal distribution. Hence, the temperature process can be modelled as a Brownian Motion.

In order to model the temperature behavior, we first consider a Vasicek process with mean reversion through the following stochastic differential equation:

$$dT_t = a(\vartheta - T_t)dt + \sigma dW_t , \quad (4)$$

where  $T_t$  is the process to be modelled,  $a \in \mathbb{R}$  is the speed of mean reversion,  $\vartheta$  the mean to which the process reverts to (constant),  $\sigma$  the volatility of the process (constant) and  $W_t$  is the Wiener process.

For the temperature process we need a  $\vartheta = \vartheta(t) = \vartheta_t$  computed according to (8),  $a = a(i) = a_i$  and  $\sigma = \sigma(i) = \sigma_i$  as functions changing over the months but constant in each month  $i$ .

Then our process becomes

$$dT_t = a_i(\vartheta_t - T_t)dt + \sigma_i dW_t . \quad (5)$$

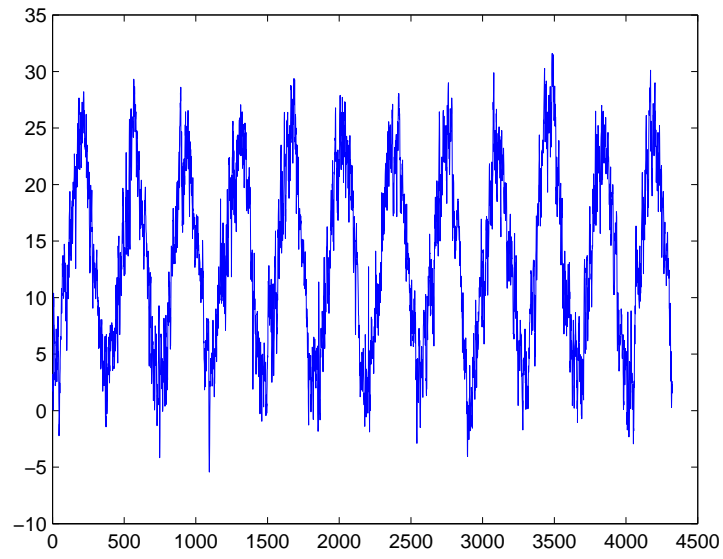


Figure 1: Daily mean temperatures at Bergamo during 1994-2005.

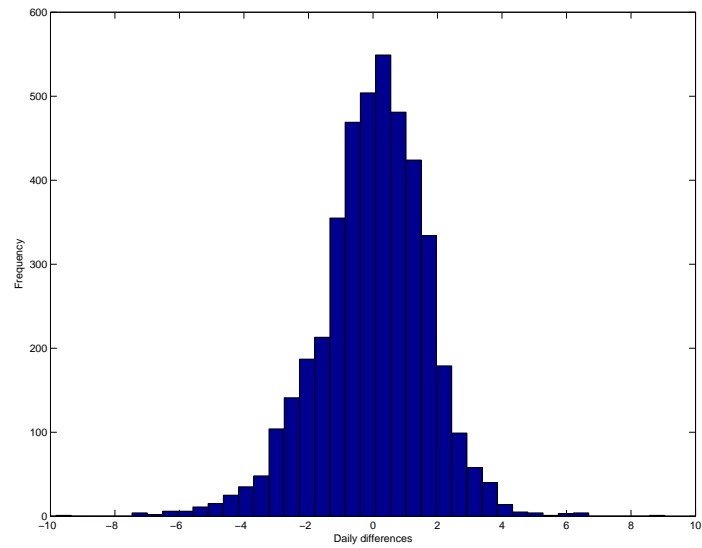


Figure 2: Daily temperature difference at Bergamo during 1994 – 2005.

Thus, we need to determine a functional form for  $\vartheta_t$  and estimates for  $a_i$  and  $\sigma_i$  from historical data. In Dornier and Queruel, (2000) [10], it is shown that the process found in (5) is not reverting to  $\vartheta_t$ ; to obtain a process that really reverts to the mean we have to add the term  $\vartheta'_t$  to the drift term in (5) so that the equation becomes

$$dT_t = \left[ a_i (\vartheta_t - T_t) + \frac{d\vartheta_t}{dt} \right] dt + \sigma_i dW_t . \quad (6)$$

The proof of reversion to the mean can be found in Appendix.

## 2.1 The mean temperature $\vartheta_t$

By observing the plot of the temperature data measured in Bergamo in the last 12 years, see Figure 1, we note a strong seasonal variation, which can be modelled by the function

$$\sin(\omega t + \varphi) , \quad (7)$$

where  $t$  is the time measured in days,  $\omega = 2\pi/365$  is the period of oscillation and  $\varphi$  is a phase angle due to the fact that the yearly minimum and maximum mean temperatures do not necessarily occur at January 1 and July 1 respectively. Moreover, the mean temperature actually increases each year (the positive trend in the data is weak but it does exist): therefore we assume a linear warming trend. A deterministic model  $\vartheta_t$  for the mean temperature at time  $t$ , is assumed to be given by

$$\vartheta_t = A + Bt + C \sin(\omega t + \varphi) , \quad (8)$$

or equivalently by

$$\vartheta_t = A + Bt + C [\cos(\varphi) \sin(\omega t) + \sin(\varphi) \cos(\omega t)] , \quad (9)$$

where we estimate the unknown parameters  $A$ ,  $B$ ,  $C$ ,  $\omega$  and  $\varphi$  so that the curve given by (9) fits the data.

In order to estimate the parameters in (9), a change of variables is operated and the constants are renamed as follows

$$\begin{cases} A = a_1 \\ B = a_2 \\ C \cos(\varphi) = a_3 \\ C \sin(\varphi) = a_4 \end{cases} \quad (10)$$

or equivalently

$$\begin{cases} A = a_1 \\ B = a_2 \\ C = \sqrt{a_3^2 + a_4^2} \\ \varphi = \arctan\left(\frac{a_4}{a_3}\right) - \pi \end{cases} \quad (11)$$

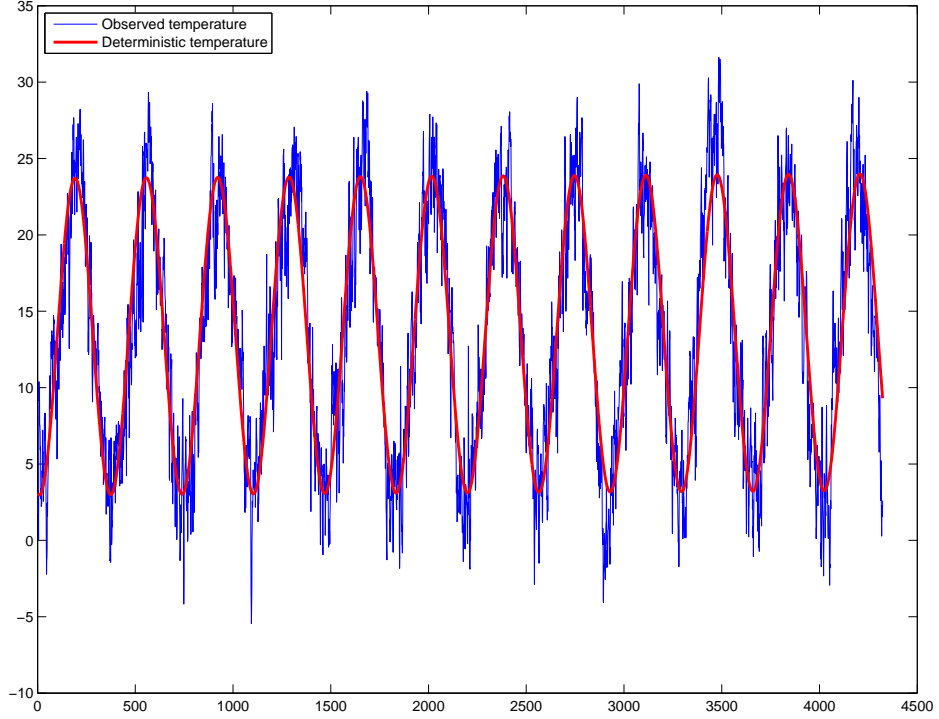


Figure 3: Comparison between measured temperatures and estimated mean  $\vartheta(t)$  at Bergamo in the years 1994 – 2005.

and we obtain

$$\vartheta_t = a_1 + a_2 t + a_3 \sin(\omega t) + a_4 \cos(\omega t) . \quad (12)$$

The numerical values of the parameters in (12) are computed by the least squares method, i.e. the parameter vector  $\xi = (a_1, a_2, a_3, a_4)$  is computed that solves

$$\min_{\xi} \|\vartheta - \mathbf{X}\|^2 , \quad (13)$$

where  $\vartheta$  is the vector whose elements are given by (12) and  $\mathbf{X}$  is the data vector. By using the series of 4323 observations of the historical daily temperatures we get

$$\begin{cases} A = 13.33 \\ B = 6.8891 \cdot 10^{-5} \\ C = 10.366 \\ \varphi = -1.7302 \end{cases} \quad (14)$$

In Figure 3 we can see a comparison between the observed temperatures and those estimated by using the deterministic approach given by  $\vartheta_t$  in the years 1994 – 2005.

## 2.2 Estimation of $\sigma_i$

For the estimation of the volatility  $\sigma_i$  we follow the same approach as in Alaton *et al.*, (2002) [1], where the quadratic variation  $\sigma_i^2$  of temperature is assumed to be different along the months in the year, but nearly constant within each month. As only the temperature mean value in each month is needed, it is not necessary to use a more elaborate model. For this reason  $\sigma_i$  is assumed to be a piece-wise constant function, with a constant value during each month. One possibility is to use an estimator based on the quadratic variation of  $T_t$  (see Basawa and Prasaka Rao, (1980) [4])

$$\hat{\sigma}_i^2 = \frac{1}{N_i} \sum_{t=0}^{N_i-1} (T_{t+1} - T_t)^2, \quad (15)$$

where  $N_i$  denotes the number of days of month  $i$  and  $t = 0$  refers to the last day of the previous month.

Another estimator is derived by discretizing (6) and using the discretized equation as a regression equation. During a given month  $i$ , the discretized equation is

$$T_t = \vartheta_t - \vartheta_{t-1} + a_i \vartheta_{t-1} + (1 - a_i) T_{t-1} + \sigma_i \epsilon_{t-1} \quad t = 1 \dots N_i, \quad (16)$$

where  $\{\epsilon_t\}_{t=1}^{N_i-1}$  are independent standard normally distributed random variables. Thus an efficient estimator of  $\sigma_i$  is (see Brockwell and Davis, (1990) [6]),

$$\hat{\sigma}_i^2 = \frac{1}{N_{i-2}} \sum_{t=1}^{N_i} (T_t - (\vartheta_t - \vartheta_{t-1}) - \hat{a}_i \vartheta_{t-1} - (1 - \hat{a}_i) T_{t-1})^2, \quad (17)$$

where  $\hat{a}_i$  is estimated in the following section. In Table 1 for each month  $i$  the estimator of  $\sigma_i$  based on the quadratic variation, the one based on the regression approach and their mean value are reported.

## 2.3 Estimation of Speed of reversion

According to Bibby and Sorensen, (1995) [5], based on observations collected during  $N_i$  days of month  $i$ , an efficient estimator  $\hat{a}_i$  of  $a_i$  is the zero of the martingale function given by

$$G(a_i) = \sum_{t=1}^{N_i} \frac{\dot{b}(T_{t-1}; a_i)}{\sigma_{i,t-1}^2} \{T_t - E[T_t | T_{t-1}]\}, \quad (18)$$

where  $\dot{b}(T_{t-1}; a_i)$  denotes the derivative with respect to  $a_i$  of the drift term

$$b(T_t, a_i) = \frac{d\vartheta_t}{dt} + a_i (\vartheta_t - T_t). \quad (19)$$

In order to obtain the solution of (18), we have to determine each of the terms  $E[T_t | T_{t-1}]$ ; thus, if we take again the process developed in (6) for a given month  $i$  and integrate between day  $(t - 1)$  and day  $t$  in month  $i$ , we find

$$T_t = \vartheta_t + e^{-a_i} (T_{t-1} - \vartheta_{t-1}) + e^{-a_i t} \int_{t-1}^t \sigma_s e^{a_i s} dW_s, \quad (20)$$



Month	Estimator 1	Estimator 2	Mean Value
January	1.6508	1.6196	1.6352
February	1.5415	1.5515	1.5465
March	1.7455	1.7209	1.7332
April	1.8480	1.8305	1.8393
May	1.8142	1.8013	1.8078
June	1.9871	1.9765	1.9818
July	1.7605	1.7298	1.7452
August	1.6305	1.6402	1.6354
September	1.4805	1.4674	1.4739
October	1.3831	1.3905	1.3868
November	1.5062	1.4933	1.4998
December	1.4912	1.4899	1.4906

Table 1: The estimators of  $\sigma_i$  based on the quadratic variation and the regression approach and their mean value.

which yields

$$E [T_t|T_{t-1}] = e^{-a_i} (T_{t-1} - \vartheta_{t-1}) + \vartheta_t , \quad (21)$$

because the expected value of an Itô integral is zero.

By substituting (21) in (18) we find

$$G_n(a_i) = \sum_{t=1}^n \frac{\dot{b}(T_{t-1}; a_i)}{\sigma_{t-1}^2} [T_t - \vartheta_t - e^{-a_i} (T_{t-1} - \vartheta_{t-1})] , \quad (22)$$

from which we obtain

$$\hat{a}_i = -\log \left( \frac{\sum_{t=1}^n \frac{\vartheta_{t-1} - T_{t-1}}{\sigma_{t-1}^2} (T_t - \vartheta_t)}{\sum_{t=1}^n \frac{\vartheta_{t-1} - T_{t-1}}{\sigma_{t-1}^2} (T_{t-1} - \vartheta_{t-1})} \right) . \quad (23)$$

Inserting the data of temperatures and the estimator  $\hat{\sigma}$  given by (15), we find the estimator  $\hat{a}_i$ . In Table 2 the values of the estimator  $\hat{a}_i$  in the twelve months are reported.

## 2.4 Generation of temperature scenarios

In this section we consider the problem of generating temperature scenarios. Using Euler approximation scheme, we discretize equation (6) obtaining

$$T_t = \vartheta_t - \vartheta_{t-1} + a_i \vartheta_{t-1} + (1 - a_i) T_{t-1} + \sigma_i \epsilon_{t-1} , \quad (24)$$

where  $\{\epsilon_t\}_{t=1}^{364}$  are independent standard normally distributed random variables. Figure 4 shows both the evolution of a simulated trajectory of the estimated temperature and its mean  $\vartheta_t$ , while Figure 5 gives the evolution of 10 scenarios of temperatures.

The following notation is used:

Month	Estimator $\hat{a}_i$
January	0.2707
February	0.2055
March	0.2017
April	0.1755
May	0.3079
June	0.2364
July	0.3051
August	0.2559
September	0.2666
October	0.1594
November	0.183
December	0.1969

Table 2: The estimator  $\hat{a}_i$  based on the formula 23.

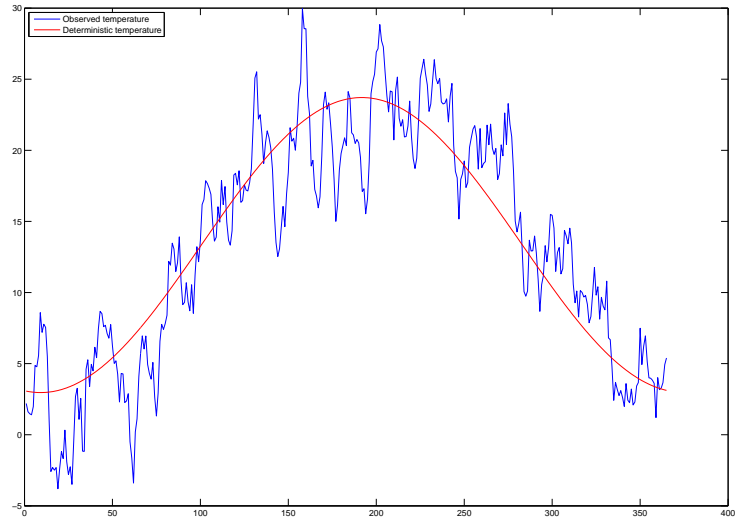


Figure 4: Simulation of sample paths of temperature and the mean estimated by Monte Carlo method.

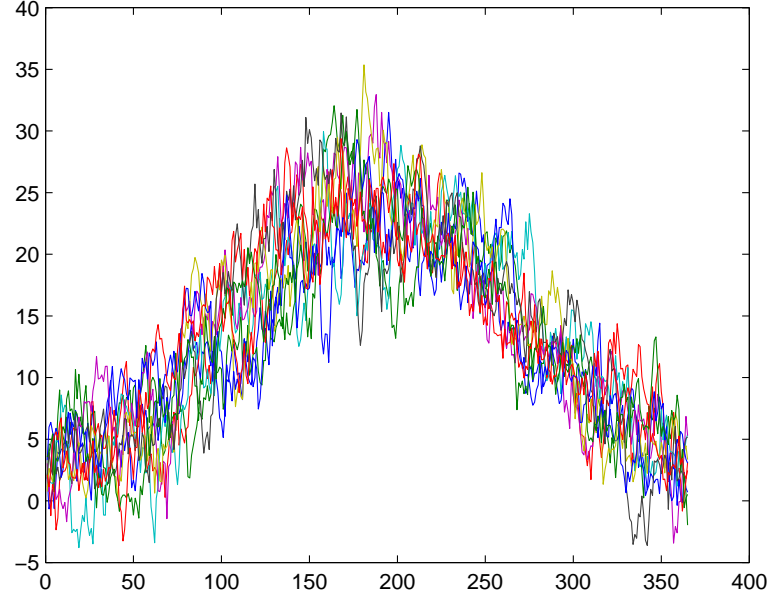


Figure 5: 10 scenarios of temperature estimated by Monte Carlo simulation.

- $\mathbf{T}^s \in \mathbb{R}^{365}$  is the vector of random variables along scenario  $s$ ,  $s = 1, \dots, N$  which we have obtained using a mean reverting process; the component  $T_t^s$  represents the daily average heating degree days for day  $t$ ,  $t = 1, \dots, 365$  along scenario  $s$ ;
- Due to the fact that the consumption data are monthly data, we generate monthly temperature scenarios from the vector  $\mathbf{T}^s$  by averaging.  $\mathbf{Tm}^s$  represents the monthly temperature scenario  $s$ , whose component  $Tm_i^s$  represents the monthly heating degree days for month  $i$ ,  $i = 1, \dots, 12$  along scenario  $s$ .
- $\bar{Tm}_i = \frac{\sum_{s=1}^N Tm_i^s}{N}$  for  $i = 1, \dots, 12$ , is the expected value over all scenarios of the random variable  $Tm_i^s$ ;
- $\Delta^s \in \mathbb{R}^{12}$  is the vector of distances of monthly heating degree days from its expected value along scenario  $s$ ,  $s = 1, \dots, N$ , i.e.  $\Delta_i^s := Tm_i^s - \bar{Tm}_i$   $i = 1, \dots, 12$ ,  $s = 1, \dots, N$ .
- $p^s$  is the probability related to each scenario  $s$ ,  $s = 1, \dots, N$ ; we assume equal probability, i.e.  $p^s = \frac{1}{N}$ ,  $s = 1, \dots, N$ ;

### 3 The stochastic OMoGaS-SV model

In the literature (see Brooks, (1981) [7], Eydeland and Wolyniec, (2003) [11], Ermoliev and Wets, (1988) [12] and Ruszczynski and Shapiro, (2003) [13]) stochastic approaches in the gas market deal mainly with the scheduling of development of gas fields, the use of gas storage and the gas delivery problem.

The stochastic version of our model, which can be classified as a two-stage stochastic program with recourse, uses the temperature  $\Delta$  as source of uncertainty. The consumptions of the first six classes of consumers are considered as dependent on temperature variations along the months.

The following notations are used:

- $I = \{i = 1, \dots, 12\}$  is the set of month indices, with  $i = 1$  corresponding to July and  $i = 12$  corresponding to the following June;
- $J = \{j = 1, \dots, 10\}$  is the set of consumer class indices;
- $\Psi = \{\psi = 1, \dots, 17\}$  is the set of energetic indices formulas;
- $S = \{s = 1, \dots, N\}$  is the set of scenario indices;
- $c_{ij}^s$  is the consumption of consumer  $j$ ,  $j = 1, \dots, 6$ , in month  $i \in I$  along scenario  $s \in S$

$$c_{ij}^s = \bar{C}_{ij} + C_{ij}\Delta_i^s, \quad j = 1, \dots, 6, i \in I, s \in S, \quad (25)$$

where  $\bar{C}_{ij}$  is the average consumption of consumer  $j$  in month  $i \in I$ ; for  $j = 7, \dots, 10$  the consumption does not depend on temperature and therefore

$$c_{ij} = \bar{C}_{ij}, \quad j = 7, \dots, 10, i \in I; \quad (26)$$

- $va_j^s$  is the annual volume of gas for consumer  $j$ ,  $j = 1, \dots, 6$ , along scenario  $s \in S$

$$va_j^s = \sum_{i=1}^{12} c_{ij}^s, \quad j = 1, \dots, 6, s \in S, \quad (27)$$

for  $j = 7, \dots, 10$  the annual volume of gas is

$$va_j = \sum_{i=1}^{12} \bar{C}_{ij}, \quad j = 7, \dots, 10; \quad (28)$$

- $vw_j^s$  is the winter volume of gas for consumer  $j$ ,  $j = 1, \dots, 6$ , along scenario  $s \in S$

$$vw_j^s = \sum_{i=5}^9 c_{ij}^s, \quad j = 1, \dots, 6, s \in S, \quad (29)$$

for  $j = 7, \dots, 10$  the winter volume of gas is

$$vw_j = \sum_{i=5}^9 \bar{C}_{ij}, \quad j = 7, \dots, 10 ; \quad (30)$$

- $r_j^s$  is the ratio of winter gas consumption with respect to the total annual consumption of consumer  $j$ ,  $j = 1, \dots, 6$ , along scenario  $s \in S$

$$r_j^s = \frac{vw_j^s}{va_j^s}, \quad j = 1, \dots, 6, \quad s \in S, \quad (31)$$

for  $j = 7, \dots, 10$  the ratio of winter gas consumption with respect to the total annual consumption is

$$r_j = \frac{vw_j}{va_j}, \quad j = 7, \dots, 10 ; \quad (32)$$

- $cd_{ij}^s$  is the peak consumption per day of customer  $j \in J$  in month  $i \in I$  for  $s \in S$

$$cd_{ij}^s = c_{ij}^s \frac{\gamma}{t_i}, \quad j \in J, \quad i \in I, \quad s \in S, \quad (33)$$

where  $t_i$  is the number of days of the month  $i \in I$  and  $\gamma$  is a parameter given by the Authority;

- $nc_j$  are the **first stage decision variables** representing the number of consumers of class  $j \in J$ , restricted to be nonnegative integers, subject to upper bounds,  $\bar{nc}_j$ ,

$$0 \leq nc_j \leq \bar{nc}_j, \quad j \in J ; \quad (34)$$

- $cm_i^s$  is the citygate consumption of month  $i \in I$  along scenario  $s \in S$

$$cm_i^s = \sum_{j=1}^6 c_{ij}^s \cdot nc_j + \sum_{j=7}^{10} c_{ij} \cdot nc_j, \quad i \in I, \quad s \in S ; \quad (35)$$

- $ca^s$  is the gas volume to be purchased for supplying the citygate consumers along scenario  $s \in S$

$$ca^s = \sum_{i=1}^{12} cm_i^s, \quad s \in S ; \quad (36)$$

- $x^s$  is the citygate loading factor along scenario  $s \in S$  and  $g$  is the **first stage decision variable** representing the maximum consumption per day above which the gas seller has to pay a penalty

$$x^s = \frac{ca^s}{365 \cdot g}, \quad s \in S ; \quad (37)$$

- $l_j$  is the loading factor of consumer class  $j$ ,  $j = 7, \dots, 10$ ;
- $s_{ki}^{+s}$ ,  $k = 0, 1, 2$  are **second stage decision variables** along scenario  $s \in S$  that represent the surplus of consumption in the peak day of winter month  $i$  ( $i = 5, \dots, 9$ ) with respect to gas availability given by the decision variable  $g$ . These variables are used in computing the penalties by  $\sum_{i=5}^9 \sum_{k=1}^2 \mu_{ki} s_{ki}^{+s}$  where  $\mu_{ki}$  is the unitary penalty in month  $i$  to be paid on the amount  $s_{ki}^{+s}$ . The unitary penalty  $\mu_{0i}$  is zero and the surplus variables  $s_{ki}^{+s}$  must satisfy the relations

$$0 \leq s_{0i}^{+s} \leq \pi_{0i} \cdot g, \quad i = 5, \dots, 9, \quad s \in S, \quad (38)$$

$$\pi_{0i} \cdot g \leq s_{1i}^{+s} \leq \pi_{1i} \cdot g, \quad i = 5, \dots, 9, \quad s \in S, \quad (39)$$

$$\pi_{2i} \cdot g \leq s_{2i}^{+s}, \quad i = 5, \dots, 9, \quad s \in S, \quad (40)$$

where  $\pi_{ki}$  represents the width of penalizations classes  $k = 0, 1$  (no upper bound for class  $k = 2$ );

- $cw^s$  is the citygate consumption in winter months along scenario  $s \in S$

$$cw^s = \sum_{i=5}^9 cm_i^s, \quad s \in S; \quad (41)$$

- $h^s$  is the ratio of winter gas consumption with respect to total annual consumption along scenario  $s \in S$

$$h^s = \frac{cw^s}{ca^s}, \quad s \in S; \quad (42)$$

- $P^s$  is the purchase price to be paid by the gas seller to the shipper along scenario  $s \in S$ : it is expressed as a linear function of  $x^s$ , and is defined as

$$P^s = QT + QS + q + m \cdot x^s, \quad s \in S; \quad (43)$$

where  $q$  is the intercept and  $m$  is the slope;  $QT$  and  $QS$  are fixed by the Italian Regulatory Authority;

- $P'_j$  is the price to be paid by the first 6 classes of consumers and is defined as

$$P'_j = (CMP + QVD) \cdot (1 - \alpha_j), \quad (44)$$

where the values of  $CMP$  and  $QVD$  are fixed by the Italian Regulatory Authority and cover raw material costs (production, importation and transport) and retail commercialization costs respectively;  $\alpha_j$  is a parameter representing possible discount fixed by the gas seller to be applied to consumer  $j$ ;

- $P_j''^s$  is the price applied by the gas seller to consumer class  $j$ ,  $j = 7, \dots, 10$  along scenario  $s \in S$

$$P_j''^s = P^s - \beta_j \cdot \left(1 - \frac{x^s}{l_j}\right) + \delta_j \cdot (r_j - h^s) + \lambda_j . \quad (45)$$

where  $\beta_j$  and  $\delta_j$  are constant values and  $\lambda_j$  is a possible recharge which can be applied to the industrial consumer class  $j$ ;

We choose as objective function the expected value of the gas seller profit:

$$\begin{aligned} w = E[ & \sum_{j=1}^6 (P_j' \cdot va_j^s \cdot nc_j) + \sum_{j=7}^{10} (P_j''^s \cdot va_j \cdot nc_j) + \\ & - P^s \cdot ca^s - \sum_{i=5}^9 \sum_{k=1}^2 \mu_{ki} s_{ki}^{+s}] . \end{aligned} \quad (46)$$

Notice that

- the expected value of revenues from the first six consumer classes is

$$E[\sum_{j=1}^6 (P_j' \cdot va_j^s \cdot nc_j)] = \sum_{j=1}^6 (P_j' \cdot \sum_{i=1}^{12} \bar{C}_{ij} \cdot nc_j) ; \quad (47)$$

- the expected value of revenues from the last four consumer classes is

$$E[\sum_{j=7}^{10} (P_j''^s \cdot va_j \cdot nc_j)] = \sum_{j=7}^{10} (nc_j E[P_j''^s]) \cdot \sum_{i=1}^{12} c_{ij} , \quad (48)$$

being the industrial consumptions independent of temperature. Notice that

$$\begin{aligned} E(P_j''^s) &= E(P^s) - \beta_j \cdot \left(1 - E\left(\frac{x^s}{l_j}\right)\right) + \delta_j \cdot (r_j - E(h^s)) + \lambda_j = \\ &= QT + QS + q - \beta_j + \delta_j r_j + \lambda_j + \left(m + \frac{\beta_j}{l_j}\right) \sum_{s=1}^N x^s p^s - \delta_j \sum_{s=1}^N h^s p^s . \end{aligned} \quad (49)$$

- the expected value of the costs is

$$E[P^s \cdot ca^s] (QT + QS + q) \sum_{i=1}^{12} E[cm_i^s] +$$

$$+\frac{m}{365 \cdot g} \left( E \left[ \sum_{i=1}^{12} (cm_i^s)^2 \right] + 2E \left[ \sum_{\substack{i,k=1 \\ k>i}}^{12} (cm_i^s) (cm_k^s) \right] \right) , \quad (50)$$

where

$$E [(cm_i^s)^2] = \sum_{s=1}^N (cm_i^s)^2 p^s ,$$

and

$$E \left[ \sum_{\substack{i,k=1 \\ k>i}}^{12} (cm_i^s) (cm_k^s) \right] = \sum_{s=1}^N \left( \sum_{\substack{i,k=1 \\ k>i}}^{12} (cm_i^s) (cm_k^s) \right) p^s ;$$

- the expected value of the penalties is

$$E \left[ \sum_{i=5}^9 \sum_{k=1}^2 \mu_{ki} s_{ki}^{+s} \right] = \sum_{s=1}^N \left( \sum_{i=5}^9 \sum_{k=0}^2 \mu_{ki} s_{ki}^{+s} \right) p^s . \quad (51)$$

The constraints of our stochastic problem are the following:

$$0 \leq nc_j \leq \overline{nc}_j , \quad j \in J , \quad (52)$$

$$\sum_{j=1}^6 cd_{ij}^s \cdot nc_j + \sum_{j=7}^{10} cd_{ij} \cdot nc_j - g \leq \sum_{k=0}^2 s_{ki}^{+s} , \quad i = 5, \dots, 9, \quad s \in S , \quad (53)$$

$$0 \leq s_{0i}^{+s} \leq \pi_{0i} \cdot g , \quad i = 5, \dots, 9, \quad s \in S , \quad (54)$$

$$\pi_{0i} \cdot g \leq s_{1i}^{+s} \leq \pi_{1i} \cdot g , \quad i = 5, \dots, 9, \quad s \in S , \quad (55)$$

$$\pi_{2i} \cdot g \leq s_{2i}^{+s} , \quad i = 5, \dots, 9, \quad s \in S , \quad (56)$$

Notice that the problem may also be formulated as a *2-stage stochastic model with recourse* as follow:

$$\max E_{\boldsymbol{\xi}} [f(x, y(\boldsymbol{\Delta}))] , \quad (57)$$

$$Ax = b , \quad (58)$$

$$T(\boldsymbol{\Delta})x + Wy(\boldsymbol{\Delta}) = h(\boldsymbol{\Delta}) , \quad (59)$$

$$x \geq 0, \quad y(\boldsymbol{\Delta}) \geq 0 , \quad (60)$$

where  $\boldsymbol{\xi} = (h(\boldsymbol{\Delta}), T(\boldsymbol{\Delta}))$  is a random vector influenced by random temperature data. In our problem the first stage decision variables  $x$  involves:



- the number of customers  $nc_j$  of class  $j \in J$ ;
- the daily capacity  $g$  above which the gas seller has to pay a penalty;

whereas the second stage decision variable  $y(\Delta)$  involves the surplus in consumption in the peak day  $s_{ki}^{+s}$  in winter month  $i$ . Furthermore the first stage constraint (58) is represented by equation (52) and the second stage constraint (59) by equations (53), (54), (55) and (56).

## 4 Results and model validations

In this section, we show the results of our stochastic model for a local gas seller who has to decide the customer portfolio structure in a village in Northern Italy (Sotto il Monte). The simulation is based on the data of thermal year 2004-2005 (for these data see Allevi *et al.*, (2005) [3]). We have developed a simulation framework based on ACCESS 97, for database management, on MATLAB release 12, for data visualization, and on GAMS release 21.5, for optimization. In the GAMS framework the DICOPT solver has been used for the nonlinear mixed integer optimization problem. DICOPT solves a series of NLP subproblems by CONOPT2 and MIP subproblems by CPLEX. The relation between the purchase price  $P^s$  and  $x^s$  is estimated by the gas seller through a linear regression using the data related to year 2004-2005 for all citygates managed by the gas seller. The regression of  $P^s$  values has also been tried on the annual volume  $ca^s$ ,  $h^s$  and  $g$  but it has been found not significant. Indeed, the value of  $R^2$ -test (see e.g. Davidson, (2000) [8]) with the regression on  $x^s$  is 0.603, therefore not highly significant. However, the introduction of non parametric regression, would introduce a more complicated function in the model. On the other side, linear regression is currently used by the gas seller in their simulations. In our case we use:

$$P^s(x^s) = QT + QS + 18.348 - 3.866 \cdot x^s, \quad (61)$$

where the intercept value 18.348 and the slope value  $-3.866$ ; the values  $QT$  and  $QS$  are given by the Italian Regulatory Authority: in our numerical experiments  $QT = 2.4953171$  Eurocent/ $Stm^3$  and  $QS = 0.63882$  Eurocent/ $Stm^3$ .

The relation between the consumption of consumer  $j$ ,  $j = 1, \dots, 6$ , in month  $i \in I$  along scenario  $s \in S$ ,  $c_{ij}^s$  and the deviation from mean value over scenarios,  $\Delta_i^s$ , is supposed to be linear with intercept equal to  $\bar{C}_{ij}$  and the other coefficient computed via a linear regression. The regression results to be significative for all the consumers.

The model has been validated by running several tests both in the deterministic (see Allevi *et al.*, (2006) [2]) and in the stochastic case. The deterministic results are reported in Table 3. For the stochastic model, we report the result obtained by solving 10000 times the problem, each time with  $N = 50$  scenarios randomly chosen with the procedure described in Section 2.4. The optimal values both in the function and in the decision variables are stable. We report in Table 4 their average over 10000 trials.

<i>Profit</i>	154265	Euro
<i>P</i>	19.67	Eurocent/Stm <sup>3</sup>
<i>ca</i>	4484406	Stm <sup>3</sup>
<i>g</i>	26399	Stm <sup>3</sup>
<i>x</i>	0.4654	

Table 3: *Optimal values for citygate Sotto il Monte in the deterministic case.*

<i>Profit</i>	152219	Euro
<i>P</i>	19.67	Eurocent/Stm <sup>3</sup>
<i>ca</i>	4484525	Stm <sup>3</sup>
<i>g</i>	26309	Stm <sup>3</sup>
<i>x</i>	0.4669	

Table 4: *Average optimal values over 10000 simulations for citygate Sotto il Monte in the stochastic case with  $N = 50$ .*

While in the deterministic case, the consumption surplus in January and February is under 10% and therefore no penalization has to be paid, in the stochastic case a nonzero penalization is applied in scenarios with high variance in consumptions. In fact the stochastic approach gives indications to the gas seller that in scenarios with colder temperatures he could face the possibility of a reduced profit due to penalties. This solution, though, allows gas seller to have the same purchase price of the deterministic case and therefore the same selling price for the industrial customer; this means that that industrial consumer is still very important and worthwhile to belong to the retail seller's portfolio.

To validate the model we analyze the sensitivity of solutions to different number of scenarios. We have run 1000 and 10000 simulations with increasing number of scenarios. In Figures 6 and 7 we report for each number of scenario the average optimal value over the corresponding number of simulations; we observe that the optimal profit converges to a value between 152200 and 152210.

## 5 Conclusions

We have proposed a stochastic model for the management of a gas sale company where the uncertainty is based on a mean reversion stochastic process for the evolution of temperature; as the number of scenarios increases, the complexity of the problem also increases: one further possibility is to devise a new algorithm that decouples computation of  $g$  from all other decision variables so that the problem becomes linear. Moreover, there exists a relation between purchase price  $p$  and international price indices, since gas seller must choose the index of reference among a certain number of admitted choices: it is possible to investigate the influence on  $P$  of future variations of these indices to help gas seller in taking his decision.

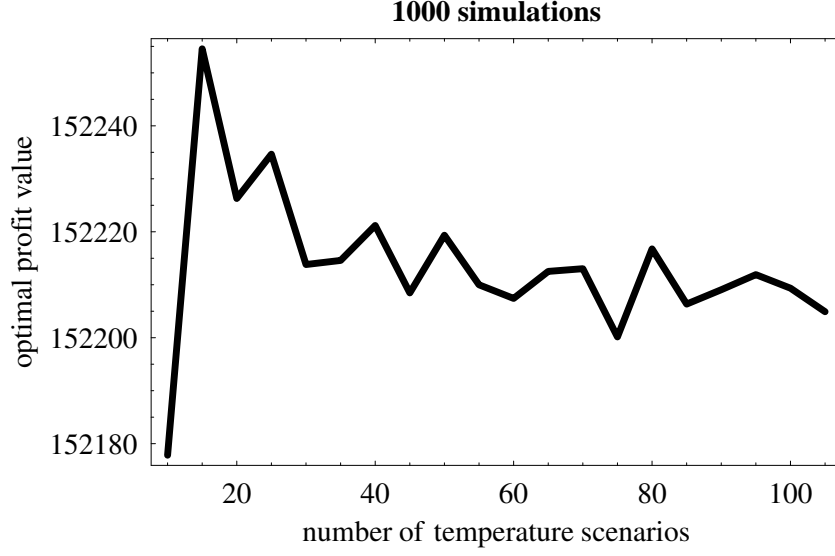


Figure 6: Case of 1000 simulations: optimal profit value as the number of scenarios increases.

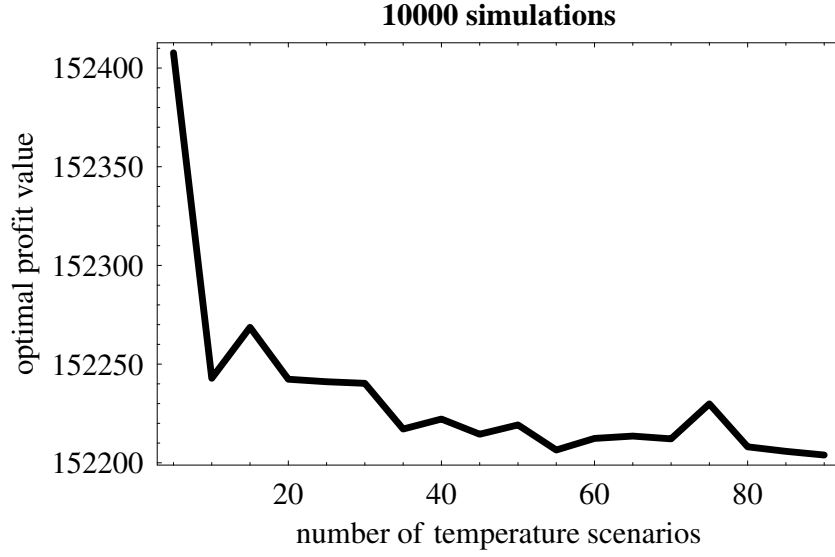


Figure 7: Case of 10000 simulations: optimal profit value as the number of scenarios increases.

## Appendix

Denoting with  $N_i$  the number of days of a specific month  $i$ , in order to prove that the process found in (6) is mean reverting we set

$$Y_t = \left[ e^{\left( \sum_{k=1}^{i-1} \int_0^{N_k} a_k ds \right) + \int_0^{t - \sum_{k=1}^{i-1} N_k} a_i ds} \right] (\vartheta_t - T_t) \quad (62)$$

$$= \left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(t - \sum_{k=1}^{i-1} N_k)} (\vartheta_t - T_t) \right],$$

then Itô's formula implies

$$\begin{aligned} dY_t = & \left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(t - \sum_{k=1}^{i-1} N_k)} \left[ \left( \frac{d\vartheta_t}{dt} + a_i(\vartheta_t - T_t) \right) dt + \right. \right. \\ & \left. \left. - \left( a_i(\vartheta_t - T_t) + \frac{d\vartheta_t}{dt} \right) dt - \sigma_t dW_t \right] \right], \end{aligned} \quad (63)$$

hence

$$Y_t - Y_0 = - \int_0^t \sigma_s \left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(s - \sum_{k=1}^{i-1} N_k)} \right] dW_s, \quad (64)$$

that is

$$\left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(t - \sum_{k=1}^{i-1} N_k)} (\vartheta_t - T_t) \right] = \vartheta_0 - T_0 - \int_0^t \left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(s - \sum_{k=1}^{i-1} N_k)} \right] \sigma_s dW_s, \quad (65)$$

but  $\vartheta_0 = T_0 = C$  and thus

$$T_t = \vartheta_t + e^{-[(\sum_{k=1}^{i-1} N_k a_k) + a_i(t - \sum_{k=1}^{i-1} N_k)]} \int_0^t \left[ e^{(\sum_{k=1}^{i-1} N_k a_k) + a_i(s - \sum_{k=1}^{i-1} N_k)} \right] \sigma_s dW_s, \quad (66)$$

from which we can see that the process reverts to its mean  $\vartheta_t$  because the expected value of an Itô Integral is zero.

## References

- [1] Alaton, P., Djehiche, B. and Stillberger, D. 2002 *On Modelling and Pricing Weather Derivatives*, Applied Mathematical Finance, Vol. 9, Issue 1.
- [2] Allevi, E., Bertocchi, M.I., Innorta, M., Vespucci, M.T. 2006 *A mixed integer nonlinear optimization model for gas sale company*, Optimization Letters DOI 10.1007/s11590-006-0012-7.
- [3] Allevi, E., Bertocchi, M.I., Innorta, M., Vespucci, M.T. 2005 *A stochastic optimization model for gas sales companies*, , Department of Mathematics, Statistic, Computer science and Applications, Working Paper 11, University of Bergamo, submitted to IMA Journal of Managements Mathematics.
- [4] Basawa, I.V. and Prakasa Rao B.L.S. 1980 *Statistical Inference for Stochastic Processes*, Academic Press.
- [5] Bibby, B. M. and Sorenses, M. 1995 *Martingale Estimation Functions for Discretely Observed Diffusion Processes*, Bernoulli vol. I numbers I/II.
- [6] Brockwell, P.J. and Davis, R.A. 1990 *Time Series: Theory and Methods*, Springer, Second edition.

- [7] Brooks, R. E. 1981 *Using generalized networks to forecast natural gas distribution and allocation during periods of shortage*, Mathematical Programming Study, 15, 23–42.
- [8] Davidson J. 2000 *Econometric Theory*, Blackwell, ISBN 0-631-21584-0.
- [9] *Deliberazione n.138, 4/12/2003, Criteri per la determinazione delle condizioni economiche di fornitura del gas naturale ai clienti finali e disposizioni in materia di tariffe per l'attività di distribuzione.*
- [10] Dornier, F. and Queruel, M. 2000 *Pricing weather derivatives by marginal value*, Quantitative Finance, V.1, Institute of Physics Publishing.
- [11] Eydeland, A., Wolyniec, K. 2003 *Energy and power risk management*, Wiley.
- [12] Ermoliev, Y., Wets, J.-B. 1988 *Numerical techniques for stochastic optimization*, Springer Verlag.
- [13] Ruszczyński, A., Shapiro, A. 2003 *Stochastic programming*, Elsevier.

**Redazione**

Dipartimento di Matematica, Statistica, Informatica ed Applicazioni  
Università degli Studi di Bergamo  
Via dei Caniana, 2  
24127 Bergamo  
Tel. 0039-035-2052536  
Fax 0039-035-2052549

La Redazione ottempera agli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche

Stampato nel 2006  
presso la Cooperativa  
Studium Bergomense a r.l.  
di Bergamo