A GAS RETAIL STOCHASTIC OPTIMIZATION MODEL BY MEAN REVERTING TEMPERATURE SCENARIOS

F. MAGGIONI[2,*], M. T. VESPUCCI[3], E. ALLEVI[1], M. I. BERTOCCHI[2] and M. INNORTA[3]

[1] *Department of Quantitative Methods, Brescia University,*
Contrada S. Chiara 50, Brescia, 25122, Italy

[2] *Department of Mathematics, Statistic, Computer Science and Applications,*
Bergamo University, Via dei Caniana 2, Bergamo, 24127, Italy
*E-mail: francesca.maggioni@unibg.it

[3] *Department of Engineering and Management, Bergamo University,*
Via Marconi 5, Dalmine, 24044, Italy

Abstract.

The paper deals with a new stochastic optimization model, named OMoGaS-SV (Optimisation 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. 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.

Keywords: Gas sale company, tariff components, mean reverting process, stochastic programming.

1. Introduction

Starting in 1999 the Italian Natural Gas market has been undergoing a libezalisation process aiming at promoting competition and efficiency, while ensuring adequate service quality standards. Timings and methods for the internal gas market liberalisation 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 2003 the Italian Regulatory Authority for Electricity and Gas,[7] 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 (Allevi *et al.*[2]), a deterministic optimization model has been developed to assist companies dealing with retail commercialization.

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 2.1. Given a weather station, let T_μ^{max} and T_μ^{min} denote the maximum and the minimum temperatures (in Celsius degrees) measured in day μ, respectively. We define the **mean temperature** of day μ as

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

Definition 2.2. Let T_μ denote the mean temperature of day μ. We define **Heating Degree Days** (HDD_μ: measure of cold in winter) and **Cooling Degree Days** (CDD_μ: measure of heat in summer) respectively as $HDD_\mu = max\left\{18 - T_\mu, 0\right\}$, $CDD_\mu = max\left\{T_\mu - 18, 0\right\}$.

For a given day HDD and CDD are the numbers of degrees of deviation from a reference temperature level in Bergamo (18° C). 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 by 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. We have plotted the daily mean temperatures at Bergamo for the 12 years; it is evident that the temperature process is mean stationary and variance stationary so it should be a mean reverting process, reverting to some cyclical function. However, because temperature process is evidently not deterministic, we must consider the presence of noise. Looking at the histogram of the daily temperature differences in Bergamo (1994–2005), it shows a good fit with the corresponding normal distribution, though the frequency of small differences in daily mean temperature is underestimated. This is the reason which brings us to choose the Brownian Motion as model of temperature process.

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

$$(2) \qquad \mathrm{d}T_t = a\left(\vartheta - T_t\right)\mathrm{d}t + \sigma\mathrm{d}W_t \ ,$$

where T_t is the process to be modelled, $a \in \mathbb{R}$ is the speed of mean reversion, ϑ the mean to which the process reverts to (constant), σ 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 (6), $a = a(i) = a_i$ and $\sigma = \sigma(i) = \sigma_i$ as functions changing over the months but constant in each month i.

Then our process becomes

$$(3) \qquad dT_t = a_i (\vartheta_t - T_t) \, dt + \sigma_i dW_t \ .$$

Thus, we need to determine a functional form for ϑ_t and estimates for a_i and σ_i from historical data. Dornier and Queruel[8] showed that the process found in (3) is not reverting to ϑ_t; to obtain a process that really reverts to the mean we have to add the term ϑ_t' to the drift term in (3) so that the equation becomes

$$(4) \qquad dT_t = \left[a_i (\vartheta_t - T_t) + \frac{d\vartheta_t}{dt} \right] dt + \sigma_i dW_t \ .$$

2.1. *The mean temperature ϑ_t*

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

$$(5) \qquad \sin(\omega t + \varphi) \ ,$$

where t is the time measured in days, $\omega = 2\pi/365$ is the period of oscillation and φ is a phase angle due to the fact that the yearly minimun 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 ϑ_t for the mean temperature at time t, is assumed to be given by

$$(6) \qquad \vartheta_t = A + Bt + C \sin(\omega t + \varphi) \ ,$$

where we estimate the unknown parameters A, B, C, ω and φ so that the curve given by (6) fits the data.

In order to estimate the parameters in (6), a change of variables is operated and the constants are renamed as follows

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

and we obtain

$$(7) \qquad \vartheta_t = a_1 + a_2 t + a_3 \sin(\omega t) + a_4 \cos(\omega t) \ .$$

The numerical values of the parameters in (7) 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

$$(8) \qquad min_\xi \|\vartheta - \mathbf{X}\|^2 \ ,$$

where ϑ is the vector whose elements are given by (7) and $\mathbf{X}$ is the data vector. By using the series of 4323 observations of the historical daily temperatures we get

$$A = 13.33, \quad B = 6.8891 \cdot 10^{-5}, \quad C = 10.366, \quad \varphi = -1.7302.$$

In Figure 1 we can see a comparison between the observed temperatures and those estimated by using the deterministic approach given by ϑ_t in the years 1994–2005.
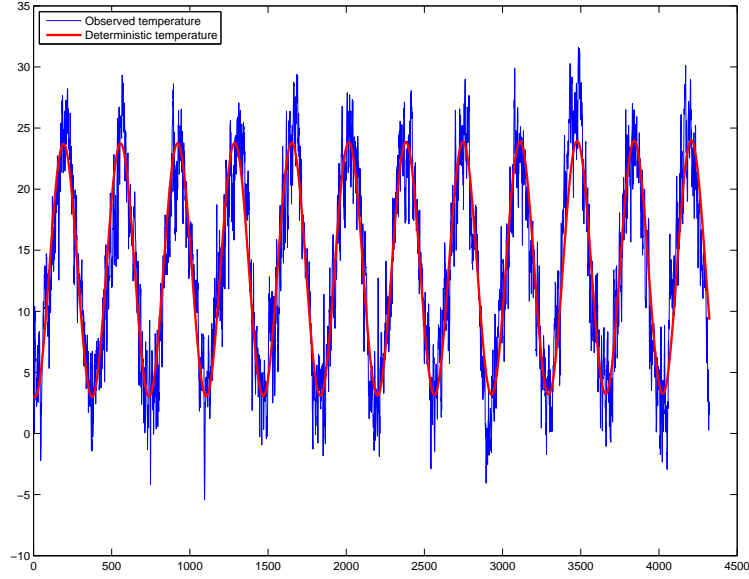
Figure 1.   Comparison between measured temperatures and estimated mean $\vartheta(t)$ at Bergamo in the years 1994–2005.

2.2. Estimation of volatility σ_i and of speed of reversion a_i

For the estimation of the volatility σ_i we follow the same approach as in Alaton *et al.*,[1] where the quadratic variation σ_i^2 of temperature is assumed to be different along the months in the year, but nearly constant within each month. For the estimation of the speed of reversion a_i we follow the same approach as in Bibby and Sorensen,[4] based on observations collected during N_i days of month i. See Maggioni *et al.*[11] for details on implementation.

2.3. Generation of temperature scenarios

In this section we consider the problem of generating temperature scenarios. Using Euler approximation scheme, we discretize equation (4) obtaining

$$(9) \qquad\qquad T_t = \vartheta_t - \vartheta_{t-1} + a_i \vartheta_{t-1} + (1 - a_i) T_{t-1} + \sigma_i \epsilon_{t-1} \ ,$$

where $\{\epsilon_t\}_{t=1}^{364}$ are independent standard normally distributed random variables. Figure 2 shows both the evolution of a simulated trajectory of the estimated temperature and its mean ϑ_t.

The following notation is used:

- $\mathbf{T}^s \in \mathbb{R}^{365}$ 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
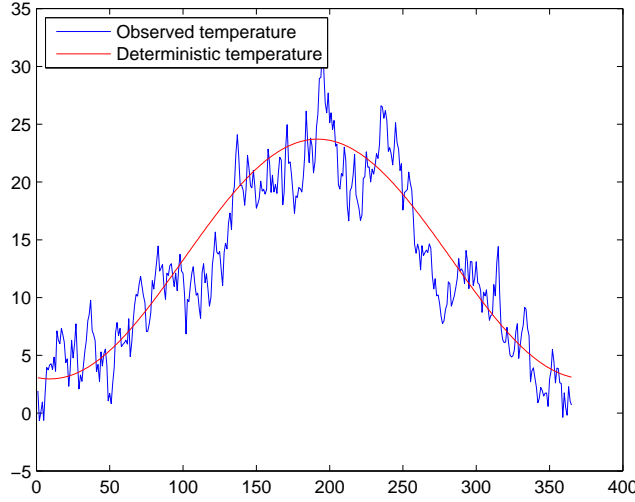
Figure 2.   Simulation of sample paths of temperature and the mean over the year 2006 estimed by using the Euler scheme.

monthly temperature scenario s, whose component Tm_i^s represents the monthly heating degree days for month i, $i = 1, \ldots 12$ along scenario s.

- $\bar{T}m_i = \dfrac{\sum_{s=1}^{N} Tm_i^s}{N}$ for $i = 1, \ldots, 12$, is the expected value over all scenarios of the random variable Tm_i^s;

- $\mathbf{\Delta}^s \in \mathbb{R}^{12}$ is the vector of distances of monthly heating degree days from its expected value along scenario s, $s = 1, \ldots, N$, i.e. $\Delta_i^s := Tm_i^s - \bar{T}m_i \qquad i = 1, \ldots, 12, \; s = 1, \ldots, N$.

- p^s is the probability related to each scenario s, $s = 1, \ldots, N$; we assume equal probability, i.e. $p^s = \dfrac{1}{N}$, $s = 1, \ldots, N$;

3. The stochastic OMoGaS-SV model

In the literature (see Brooks,[5] Eydeland and Wolyniec,[9] Ermoliev and Wets[10] and Ruszczynski and Shapiro[12]) 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 $\mathbf{\Delta}$ 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:

- c_{ij}^s is the consumption of consumer j, $j = 1, \ldots, 6$, in month i, $i = 1, \ldots, 12$ along scenario s, $s = 1, \ldots, N$: $c_{ij}^s = \bar{C}_{ij} + C_{ij}\Delta_i^s$, where $\bar{C}_{ij}$ is the average consumption of consumer j in month i; for $j = 7, \ldots, 10$ the consumption does not depend on temperature and therefore $c_{ij} = \bar{C}_{ij}$;

- va_j^s is the annual volume of gas for consumer $j = 1, \ldots, 6$, along scenario s, $s = 1, \ldots, N$: $va_j^s = \sum_{i=1}^{12} c_{ij}^s$; for $j = 7, \ldots, 10$ the annual volume of gas is $va_j = \sum_{i=1}^{12} \bar{C}_{ij}$;

- vw_j^s is the winter volume of gas for consumer j, $j = 1, \ldots, 6$, along scenario s, $s = 1, \ldots, N$: $vw_j^s = \sum_{i=5}^{9} c_{ij}^s$; for $j = 7, \ldots, 10$ the winter volume of gas is $vw_j = \sum_{i=5}^{9} \bar{C}_{ij}$;

- r_j^s is the ratio of winter gas consumption with respect to the total annual consumption of consumer j, $j = 1, \ldots, 6$, along scenario s, $s = 1, \ldots, N$: $r_j^s = \frac{vw_j^s}{va_j^s}$; for $j = 7, \ldots, 10$ the ratio of winter gas consumption with respect to the total annual consumption is $r_j = \frac{vw_j}{va_j}$;

- cd_{ij}^s is the peak consumption per day of customer j, $j = 1, \ldots, 10$, in month i, $i = 1, \ldots, 12$ for $s = 1, \ldots, N$: $cd_{ij}^s = c_{ij}^s \frac{\gamma}{N_i}$, where N_i is the number of days of the month i and γ is a parameter given by the Authority;

- cm_i^s is the citygate consumption of month i, $i = 1, \ldots, 12$ along scenario s, $s = 1, \ldots, N$: $cm_i^s = \sum_{j=1}^{6} c_{ij}^s \cdot nc_j + \sum_{j=7}^{10} c_{ij} \cdot nc_j, \quad i = 1, \ldots, 12$;

- nc_j are the **first stage decision variables** representing the number of consumers of class j, restricted to be nonnegative integers, subject to upper bounds, $\overline{nc}_j$, $0 \le nc_j \le \overline{nc}_j \quad j = 1, \ldots, 10$;

- ca^s is the gas volume to be purchased for supplying the citygate consumers along scenario s, $s = 1, \ldots, N$: $ca^s = \sum_{i=1}^{12} cm_i^s$;

- x^s is the citygate loading factor along scenario s, $s = 1, \ldots, N$ and g is the **first stage decision variable** representing the maximum consumption per day above which the gas seller has to pay a penalty: $x^s = \frac{ca^s}{365 \cdot g}$;

- l_j is the loading factor of consumer class j, $j = 7, \ldots, 10$;

- $s_{ki}^{+\,s}$, $k = 0, 1, 2$ are **second stage decision variables** along scenario s, $s = 1, \ldots, N$ that represent the surplus of consumption in the peak day of winter month i ($i = 5, \ldots, 9$) with respect to gas availability given by the first stage 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 μ_{ki} is the unitary penalty in month i to be paid on the amount $s_{ki}^{+\,s}$. The unitary penalty μ_{0i} is zero and the surplus variables $s_{ki}^{+\,s}$ must satisfy the relations $0 \le s_{0i}^{+\,s} \le \pi_{0i} \cdot g$, $\pi_{0i} \cdot g \le s_{1i}^{+\,s} \le \pi_{1i} \cdot g$, $\pi_{2i} \cdot g \le s_{2i}^{+\,s}$, $i = 5, \ldots, 9$, , $s = 1, \ldots, N$, where π_{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, $s = 1, \ldots, N$: $cw^s = \sum_{i=5}^{9} cm_i^s$;

- h^s is the ratio of winter gas consumption with respect to total annual consumption along scenario s, $s = 1, \ldots, N$: $h^s = \frac{cw^s}{ca^s}$;

- P^s is the purchase price to be paid by the gas seller to the shipper along scenario s, $s = 1, \ldots, N$: it is expressed as a linear function of x^s, and is defined as $P^s = QT + QS + q + m \cdot x^s$; 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 $P_j' = (CMP + QVD) \cdot (1 - \alpha_j)$, 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; α_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 $j = 7, \ldots, 10$ along scenario s, $s = 1, \ldots, N$

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

where β_j and δ_j are constant values and λ_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:

$$w = E[\sum_{j=1}^{6}(P_j' \cdot va_j^s \cdot nc_j) + \sum_{j=7}^{10}(P_j''^{\,s} \cdot va_j \cdot nc_j)+$$

$$(1) \qquad -P^s \cdot ca^s - \sum_{i=5}^{9}\sum_{k=1}^{2}\mu_{ki}s_{ki}^{+\,s}] \ .$$

For details on implementation of the objective function see Maggioni *et al.*.[11]
The constraints of our stochastic problem are the following:

$$(2) \qquad 0 \leq nc_j \leq \overline{nc}_j \qquad j = 1,\ldots,10 \ ,$$

$$(3) \qquad \sum_{j=1}^{6} cd_{ij}^s \cdot nc_j + \sum_{j=7}^{10} cd_{ij} \cdot nc_j - g \leq \sum_{k=0}^{2} s_{ki}^{+^s} \qquad i = 5,\ldots,9, \ s = 1,\ldots,N \ ,$$

$$(4) \qquad 0 \leq s_{0i}^{+^s} \leq \pi_{0i} \cdot g \ , \qquad i = 5,\ldots,9, \ s = 1,\ldots,N \ ,$$

$$(5) \qquad \pi_{0i} \cdot g \leq s_{1i}^{+^s} \leq \pi_{1i} \cdot g \ , \qquad i = 5,\ldots,9, \ s = 1,\ldots,N \ ,$$

$$(6) \qquad \pi_{2i} \cdot g \leq s_{2i}^{+^s} \ , \qquad i = 5,\ldots,9, \ s = 1,\ldots,N \ .$$

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.*[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[6]) 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 $P^s(x^s) = QT + QS + 18.348 - 3.866 \cdot x^s$, 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 c_{ij}^s of consumer j, $j = 1, \ldots, 6$, in month $i = 1, \ldots, 12$ along scenario $s = 1, \ldots, N$ and the deviation from mean value over scenarios, Δ_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 $al.$[2]) and in the stochastic case, see Table 1. 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 2.3. The optimal values both in the function and in the decision variables are stable. We report in the second column of Table 1 their average over 10000 trials.

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 have analyzed the sensitivity of solutions to different number of scenarios. We have run 1000 and 10000 simulations with increasing number of scenarios. In Figures 1 and 2 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.
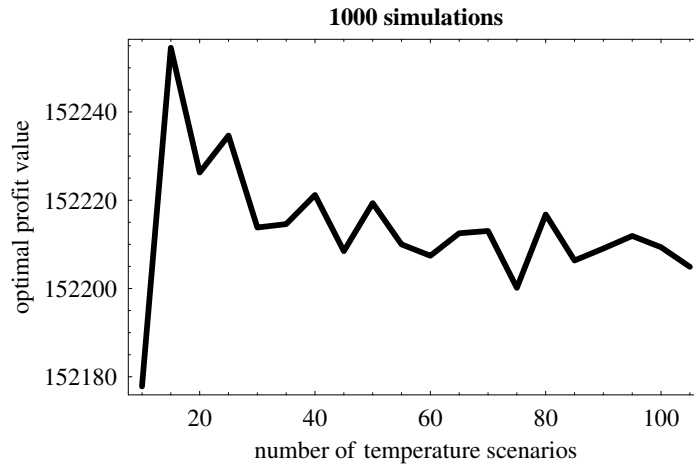

Figure 1.   Case of 1000 simulations: optimal profit value as the number of scenarios increases.
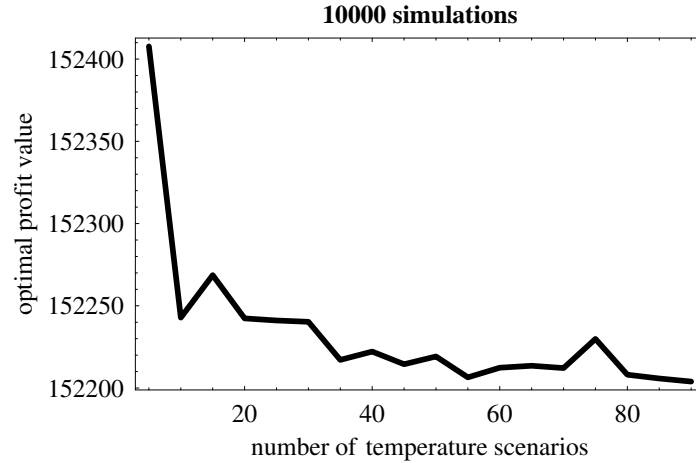
Figure 2. Case of 10000 simulations: optimal profit value as the number of scenarios increases.

Table 1. Optimal values for citygate Sotto il Monte in the deterministic and stochastic case.

	Deterministic	Stochastic	
$Profit$	154265	152219	Euro
P	19.67	19.67	Eurocent/Stm^3
ca	4484406	4484525	Stm^3
g	26399	26309	Stm^3
x	0.4654	0.4669	

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.

REFERENCES

1. P. Alaton, B. Djehiche and D. Stillberger, *Applied Mathematical Finance*, **9**, 1 (2002).
2. E. Allevi, M.I. Bertocchi, M. Innorta, M.T. Vespucci, A mixed integer nonlinear optimization model for gas sale company, *Optimization Letters*, **1**(1), 61 Springer (2007).
3. E. Allevi, M.I. Bertocchi, M. Innorta, M.T. Vespucci, A stochastic optimization model for a gas sale company, IMA Journal of Management Mathematics, 1-14, DOI:10.1093/imaman/dpm004 (2007).
4. B.M. Bibby, and M. Sorensen, *Bernoulli*, **1**(I/II), 17 (1995).
5. R.E. Brooks, *Mathematical Programming Study*, **15**, 23 (1981).
6. Davidson J. 2000 Econometric Theory, Blackwell, ISBN 0-631-21584-0.

7. *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.*

8. F. Dornier, and M. Queruel, *Caution to the Wind*, Weather Risk Special Report, Energy & Power Risk Management/Risk Magazine (2000).

9. A. Eydeland and K. Wolyniec, *Energy and power risk management*, Wiley (2003).

10. Y. Ermoliev and J.-B. Wets, *Numerical techniques for stochastic optimization*, Springer Verlag, (1988).

11. F. Maggioni, M.T. Vespucci, E. Allevi, M.I. Bertocchi, M. Innorta, A stochastic optimization model for a gas sale company using mean reverting process modeling of the temperature, *Department of Mathematics, Statistic, Computer science and Applications*, Working Paper **4**, University of Bergamo, (2006).

12. A. Ruszczynski and A. Shapiro, *Stochastic programming*, Elsevier, (2003).