Modeling the price of natural gas with temperature and oil price as exogenous factors

Jan Müller[a,*], Guido Hirsch[b], Alfred Müller[a]

[a] *Department Mathematik, Universität Siegen, Walter-Flex-Str. 3, 57072 Siegen, Germany*
[b] *EnBW Trading GmbH, Durlacher Allee 93, 76131 Karlsruhe, Germany*

Abstract

The literature on stochastic models for the spot market of gas is dominated by purely stochastic approaches. In contrast to these models, Stoll and Wiebauer (2010) propose a fundamental model with temperature as an exogenous factor. In Central European markets there is another important fundamental driver of the gas price, namely the oil price. This is due to import contracts with Russia and Norway, where the import price is fixed by so-called oil price formulas. In this paper the model of Stoll and Wiebauer (2010) is extended by an oil price component. This component is an approximation of the unknown oil price formulas of the import contracts. It is shown that this new model can explain the price movements of the last few years much better than previous models.

1. Introduction

During the last years trading of natural gas has become more important. The traded quantities over-the-counter and on energy exchanges have strongly increased and new products have been developed. For example, swing options increase the flexibility of suppliers and they are used as an instrument for risk management purposes. Important facilities for the security of supply are gas storages. The storages are filled in times of low consumption and emptied in times of high consumption which usually coincides with

*Corresponding author
Email addresses: jan.mueller@uni-siegen.de (Jan Müller), gu.hirsch@enbw.com (Guido Hirsch), mueller@mathematik.uni-siegen.de (Alfred Müller)

Preprint submitted to Energy & Finance 2012 *September 30, 2012*

low and high prices of natural gas. Although the liquidity is comparably low, a market for gas storages exists. This means, physical gas storages or gas storage contracts are traded.

These two examples of complex options illustrate the need of pricing methods. Due to the complexity analytic pricing formulas do not exist. Therefore models for the use in Monte Carlo simulations are needed for valuation purposes. As both options (partly) rely on the spot market we need a stochastic price model for the daily prices at the spot market generating adequate gas price scenarios.

The literature on stochastic gas price models is dominated by purely stochastic approaches. The one and two factor models by Schwartz (1997) and Schwartz and Smith (2000) are general approaches applicable to many commodities, such as oil and gas. Cortazar and Schwartz (2003) present a three factor model for the term structure of oil prices. These models can be applied to gas prices as well. The various factors represent short and long term influences on the price.

Extensions of these factor models are given by Jaillet, Ronn and Tompaidis (2004) and Xu (2004). Especially the inclusion of deterministic functions to cover the seasonalities within gas prices is considered. Cartea and Williams (2008) introduce a two factor model including a function for the seasonality. Their focus is on the market price of risk. An important application of gas price models is the valuation of gas storage facilities. Within this context, Chen and Forsyth (2006) and Boogert and de Jong (2011) propose gas price models. Chen and Forsyth (2006) analyze regime-switching approaches incorporating mean-reverting processes and random walks. The class of factor models is extended by Boogert and de Jong (2011). The three factors in their model represent short and long term fluctuations as well as the behavior of the summer-winter spread.

In contrast to these models Stoll and Wiebauer (2010) propose a fundamental model with temperature as an exogenous factor. They use the temperature component as an approximation of the filling level of gas storages which have a remarkable influence on the price. In this paper we will extend the model of Stoll and Wiebauer (2010) by introducing another exogenous factor to their model: the oil price. Our oil price component approximates the influence of gas import contracts indexed by oil price formulas.

The rest of the paper is organized as follows. In chapter 2 we will introduce the model by Stoll and Wiebauer (2010) including a short description of their model for the temperature component. In their model the influence of the temperature is described by so called normalized cumulated heating degree days. Then we describe our new oil price component in chapter 3. We discuss the different oil price formulas used in the market, and how we have chosen an approximation of it in our model. After introducing the stochastic processes that we use to model oil prices and temperature, we finally fit the model to data in chapter 4. It turns out that for the innovations of the process a heavy tailed distribution like NIG is more appropriate than the classical normal distribution. We finish with a short conclusion in chapter 5.

2. The model by Stoll and Wiebauer (2010)

Modeling the price of natural gas in Central Europe requires knowledge about the structure of supply and demand. On the supply side there are only a few sources in Central Europe. Most of the natural gas needs to be imported from Norway and Russia. On the demand side there are mainly three groups of gas consumers: Households, industrial companies and gas fired power plants. While households only use gas for heating purposes at low temperatures, industrial companies use gas as heating and process gas. Households and industrial companies are responsible for about 90 percent of total gas demand.

These two groups of consumers cause seasonalities in the gas price:

- Weekly seasonality: Many industrial companies do not need gas on weekends. Their operation is restricted to working days.

- Yearly seasonality: Heating gas is needed in winter when temperatures are low.

An adequate gas price model has to incorporate these seasonalities as well as stochastic deviations of these.

Stoll and Wiebauer (2010) propose a model meeting these requirements and incorporating another major influence factor: the temperature. Somehow the temperature dependency is already covered by the deterministic yearly seasonality. The lower the temperature, the higher the price. But the temperature influence is more complex than this. A day with average temperature of zero degrees at the end of a long and cold winter has a different impact on the price than a daily average of zero at the end of a "warm" winter. Similarly, a cold day at the end of a winter has a different impact on the price than a cold day at the beginning of the winter.

The different impacts are due to gas storages which are essential to cover the demand in winter. The total demand for gas is higher than the capacities of the gas pipelines from Norway and Russia. Therefore gas provider use gas storages. These storages are filled during summer (at low prices) and emptied in winter months. At the end of a long and cold winter most gas storages will be almost empty. Therefore additional cold days will lead to comparatively higher prices than in a normal winter.

The filling level of all gas storages in the market would be the adequate variable to model the gas price. Unfortunately, there is no data covering this information. Therefore we need a different variable describing the same situation. As the filling levels of gas storages are strongly related to the demand for gas which in turn depends on the temperature, an adequate variable can be derived from the temperature.

Stoll and Wiebauer (2010) use **normalized cumulated heating degree days** to cover the influence of temperature on the gas price. They define a temperature of 15 °C as the limit of heating. Any temperature below 15 °C makes households and companies switch on their heating systems. Heating degree days are measured by $HDD_t = \max(15 - T_t, 0)$ where T_t is the average temperature of day t. As mentioned above the impact on the price depends on the number of cold days observed so far in the winter. Cumulation of heating degree days over a winter leads to a number indicating how cold the winter was. In this context we refer to winter as the 1st of October and the 181 following days till end of March. We will write $HDD_{d,w}$ for HDD_t, if t is day number d of winter w. Then we can define the cumulated heating degree days on the day d in winter w as

$$CHDD_{d,w} = \sum_{k=1}^{d} HDD_{k,w} \text{ for } 1 \leq d \leq 182.$$

The impact of cumulated heating degree days on the price depends on the comparison to a normal winter. This information is included in normalized cumulated heating degree days

$$\Lambda_{d,w} = CHDD_{d,w} - \frac{1}{w-1} \sum_{\ell=1}^{w-1} CHDD_{d,\ell} \text{ for } 1 \leq d \leq 182.$$

We will use Λ_t instead of $\Lambda_{d,w}$ for simplicity, if t is a day in a winter. The definition of Λ_t for a summer day is described by a linear return to zero during summer. Although there might be cold days between April and September this time of the year is usually used to refill gas storages. The impact of cold days on the price decreases due to increasing filling levels. We take account for this situation by the linear part of normalized cumulated heating degree days (see figure 1). Positive values of $\Lambda_{d,w}$ describe winters colder than the average. Λ_t is included into the gas price model by a regression approach. As

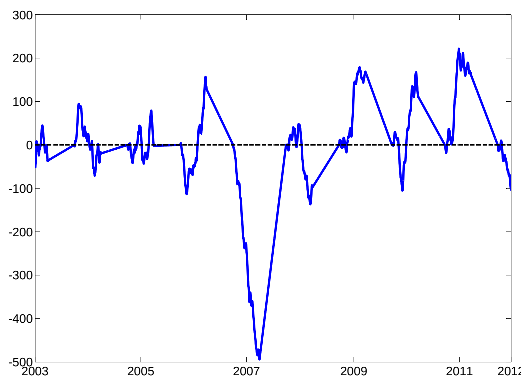

Figure 1: Normalized cumulated heating degree days in Düsseldorf, Germany, for 2003-2011.

the seasonal components and the normalized cumulated heating degree days are linear with respect to the parameters we can use ordinary least squares regression for parameter estimation. The complete model can be written as

$$G_t = m_t + \alpha \cdot \Lambda_t + X_t^{(G)} + Y_t^{(G)} \tag{1}$$

with the day-ahead price of gas G_t, the deterministic seasonality m_t, the normalized cumulated heating degree days Λ_t, an ARMA process $X_t^{(G)}$ and a geometric Brownian motion $Y_t^{(G)}$. For model calibration day-ahead gas prices from TTF and daily average temperatures from Düsseldorf, Germany, are used. The fit to historical prices where outliers have been removed (see section 4 for details on treatment of outliers) can be seen in figure 2.

3. The oil price dependence of gas prices

The model described in (1) is capable to cover all influences on the gas price related to changes in temperature. But changes for economic reasons

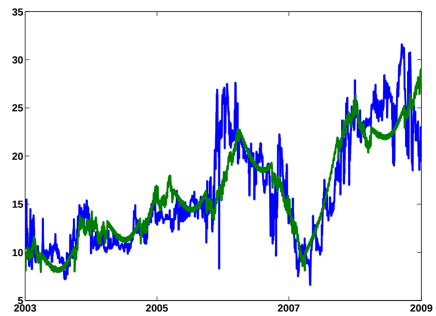

Figure 2: The model in (1) fitted to TTF prices from 2003-2009.

are not covered by that model. This was observable in the economic crisis 2008/2009 (see figure 5). During the crisis the demand for gas by industrial companies decreased by more than 10 percent. As gas is imported by long term oil price indexed supply contracts with **take-or-pay clauses** the gas importers had to take large quantities of gas although their customers had a low demand. These surplus quantities were sold at low prices on the day-ahead market. In this period of time the temperature was not relevant for the price as there was enough gas anyways.

The low prices on the day-ahead market were the result of high quantities that had to be sold in times of low demand. As the oil price also was low due to the crisis resulting in a low demand of oil, one could use the oil price as an exogenous factor explaining the gas price. If we want to incorporate the link between oil and gas in the model we need to know how the price indexation in import contracts works. The import price of gas usually is an average of past oil prices. The pricing in import contracts is done via **oil price formulas**. These formulas consist of three parameters:

1. The number of averaging months. The gas price is the average of past oil prices within a certain number of months.
2. The time lag. Possibly there is a time lag between the months the average is taken of and the months the price is valid for.
3. The number of validity months. The price is valid for a certain number of months.

An example of a 3-1-3 formula is given in figure 3.

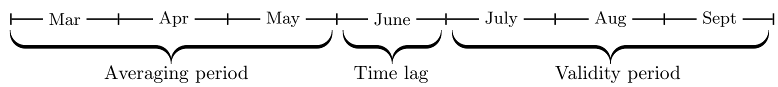

Figure 3: In a 3-1-3 formula the price is determined by the average price of 3 months (March to May). This price is valid for July to September. The next day of price fixing is the 1st of October.

The formulas used in the import contracts are not known for all market participants. Therefore we have to do some sort of estimation for our model. Theoretically any choice of three natural numbers is possible. But from other products, like oil indexed swing options, we know that some formulas are more popular than others. Examples of common formulas are 3-1-1, 3-1-3, 6-1-1, 6-1-3 and 6-3-3.

As there are many different import contracts with possibly different oil price formulas we cannot ensure that one of the mentioned formulas is able to explain the price behavior. The mixture of different formulas might affect the price in the same way as one of the common formulas or a similar one. Therefore we compare the different formulas in a regression model. Strict application of the formula means that we have jumps in the price at each day of price fixing. The impact on the gas price will be more smooth, however. The new price determined on a fixing day is the result of averaging a number of past oil prices. The closer to the fixing day the more prices for the averaging are known. Therefore market participants have reliable estimations of the new import price. If the new price will be higher it is cheaper to buy gas in advance and store it. This increases the day-ahead price prior to the fixing day and leads to a smooth transition from the old to the new price level on the day-ahead market.

This behavior of market participants leads to some smoothness of the price. In order to include this fact in our model we use a smoothed oil price formula. A sophisticated smoothing approach for forward price curves is introduced by Benth, Koekkebakker and Ollmar (2007). They claim some smoothness conditions on the borders between different price intervals. It is shown that splines of order four meet all these requirements and make sure that the result is a smooth curve. As oil price formulas are step functions like forward price curves this approach is applicable to our situation.

If the number of validity months is equal to 1 we use a moving average instead of a (smoothed) step function to simplify matters. This approach takes account for increasing information about the future price. Both approaches lead to a significant improvement of the model fit. Regarding the goodness of fit there is no remarkable difference between both approaches which justifies the use of the simpler method (see figure 4).

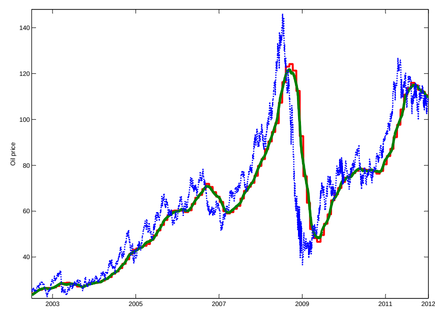

Figure 4: The price of oil (blue), the 6-0-1 oil price formula (red) and the moving average of 180 days (green).

4. Model calibration with temperature and oil price

After justification of the oil price component as a fundamental factor for our model we need to choose a way to include it in our existing model and we need to decide about the oil price formula to be used. So far we took care to keep our model linear in all parameters, i.e. that we can use ordinary least squares regression methods for the parameter estimation. Thus, we include the oil price component in a linear way as well.

For the choice of the best oil price formula we use the R^2 as the measure of goodness of fit. We choose the reasonable oil price formula leading to the highest value of R^2. Reasonable in this context means, that we restrict our analysis to formulas that are equal or similar to the ones known from other oil price indexed products (compare section 3). The result of this comparison is a 6-0-1 formula (see figure 6). Although this is not a common formula there is an explanation for it: The gas price decreased approximately six months lather than the oil price in the crisis. This major price movement needs to be covered by the oil price component. As explained above we replace the step function by a moving average. Taking the moving average of 180 days is a good approximation of the 6-0-1 formula. All in all the oil price component increases the R^2 as our measure of goodness of fit from 0.35 to 0.83 (see figure 5).

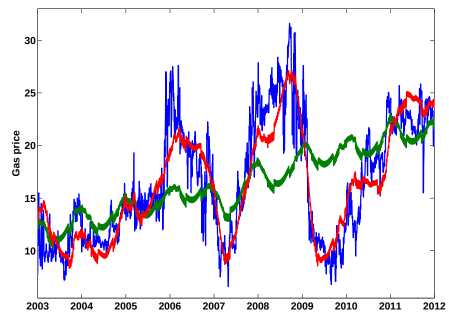

Figure 5: The model of Stoll and Wiebauer (2010) (green line) and our model (red line) fitted to historical gas prices (blue line).

Therefore we model the gas price by

$$G_t = m_t + \alpha_1 \Lambda_t + \alpha_2 \Psi_t + X_t^{(G)} \tag{2}$$

with Ψ_t being the oil price formula.

For parameter estimation of our model we use day-ahead gas prices from TTF. The trade on TTF has a longer history of high trading volumes than the neighboring markets. Data from 2003-2011 is used for this model. Temperature data from Düsseldorf, Germany, is available from 1969-2011. For the estimation of the oil price component we use prices of Brent traded on the IntercontinentalExchange (ICE). The data is available with a longer history than gas prices. We use the data from 2002-2011. Using this data we can estimate all parameters applying ordinary least squares regression after some outliers are removed from the gas price data, G_t.

Due to e.g. technical problems or a fire at a major gas storage the gas price deviated from its normal price level which was determined by temperature and oil price formulas. Thus, we exclude the prices on these occasions by an outlier treatment proposed by Weron (2006). Values outside a range around a running median are declared to be outliers. The range is defined by a multiple of the standard deviation. The identified outliers are replaced by an average of neighboring values.

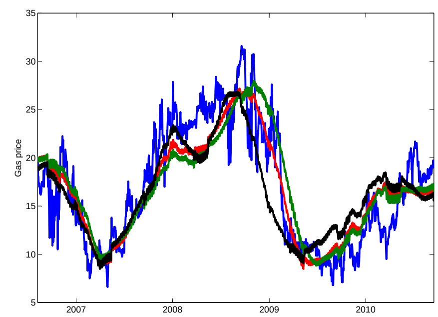

Figure 6: Comparison of different oil price components in the model: 6-0-1 formula (red line), 6-1-1 formula (green line) and 3-0-1 formula (black line) fitted to the historical prices (blue line).

Altogether these model components give fundamental explanations for the historical day-ahead price behavior. Short term deviations are included by a stochastic process (see 4.3). Long term uncertainty due to the uncertain development of the oil price is included by the oil price process. Therefore our model is able to generate reasonable scenarios for the future (see figure 7). We will specify the stochastic models for the exogenous factors Ψ_t and

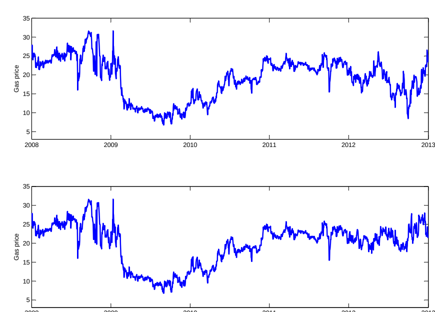

Figure 7: The historical gas price (2008-2012) and two realizations of the gas price process for 2012-2013.

Λ_t as well as the stochastic process $X_t^{(G)}$ in the following.

4.1. Oil price model

As the time series of oil prices does not contain any seasonalities, we model the oil price without any deterministic function or fundamental factors. Instead we apply the two factor model by Schwartz and Smith (2000). They divide the log price into two factors: one for short term variations and one for long term dynamics.

$$\psi_t = \exp\left(\chi_t + \xi_t\right)$$

with an AR(1) process χ_t and a Brownian motion ξ_t. These processes are correlated. Using price data of Brent Crude future contracts from the ICE we can apply the Kalman filter to estimate the model parameters. The process (ψ_t) is used to derive the process (Ψ_t) in (2).

4.2. Temperature model

When modeling daily average temperature we can make use of a long history of temperature data. Here a yearly seasonality and a linear trend can be identified. Therefore we use a temperature model closely related to the one proposed by Benth and Benth (2007).

$$T_t = a_1 + a_2 t + a_3 \sin\left(\frac{2\pi t}{365.25}\right) + a_4 \cos\left(\frac{2\pi t}{365.25}\right) + X_t^{(T)} \tag{3}$$

with $X_t^{(T)}$ being an AR(3) process. The model fit with respect to the deterministic part (ordinary least squares regression) and the AR(3) process is shown in figure 8. The process (T_t) is then used to define the derived process (Λ_t) of normalized cumulated heating degree days as described in 2.

4.3. Stochastic process

The fit of normalized cumulated heating degree days, oil price formula and deterministic components to the gas price via ordinary least squares regression results in a residual time series. These residuals contain all unexplained, "random" deviations from the usual price behavior.

The residuals exhibit a strong autocorrelation to the first lag. Therefore an AR(1) process provides a good fit. The empirical innovations of the process show more heavy tails than a normal distribution (compare Stoll and Wiebauer (2010)). Therefore we apply a distribution with heavy tails. The

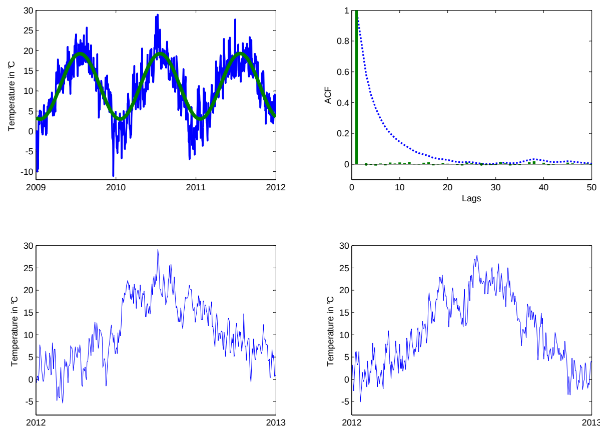

Figure 8: Top left: Fit of deterministic function (green line) to the historical daily average temperature (blue) in Düsseldorf, Germany. Top right: Autocorrelation function of residual time series (dotted) and innovations of AR(3) process (stems). Bottom: Two realizations of the temperature process.

normal-inverse gaussian (NIG) distribution remarkably increases the goodness of fit (see figure 9). Recall that a random variable X is NIG-distributed if there is a representation

$$X \stackrel{d}{=} \mu + \beta Y + \sqrt{Y} Z$$

with $Z \sim \mathcal{N}(0,1)$ and $Y \sim N^{-}(-1/2, \delta^2, \alpha^2 - \beta^2)$, the inverse Gaussian distribution as a special case of the generalized inverse Gaussian distribution. The class of generalized hyperbolic distributions including the NIG distribution was introduced by Barndorff-Nielsen (1978).

Both the distribution of the innovations and the parameters of autoregressive processes are estimated using maximum likelihood estimation.

5. Conclusion

The spot price model by Stoll and Wiebauer (2010) with only temperature as an exogenous factor is not able to explain the gas price behavior during the last years. We have shown that the extension by another exogenous factor remarkably improves the model fit on the history. This factor, the oil price

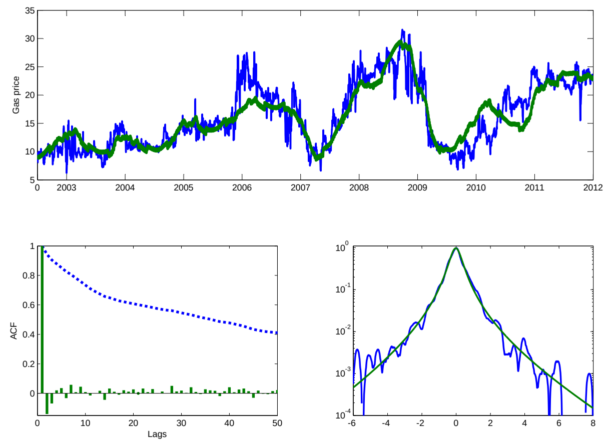

Figure 9: Top: Fit of deterministic function and exogenous components (green line) to the historical gas price (blue). Bottom left: ACF of residual time series (dotted) and innovations of AR(1) process (stems). Bottom right: Fit of NIG distribution (green) to kernel density of empirical innovations (blue).

component, approximates the oil price formulas in gas import contracts. This fundamental reason and the improvement of model fit give justification for the inclusion of the model component. The resulting simulation paths from the model are reliable and lead to reasonable valuation results.

Barndorff-Nielsen, O.E., 1978. Hyperbolic distributions and distributions on hyperbolae. Scandinavian Journal of Statistics 5, 151–157.

Benth, F.E., Benth, J.S., 2007. The volatility of temperature and pricing of weather derivatives. Quantitative Finance 7, 553–561.

Benth, F.E., Koekkebakker, S., Ollmar, F., 2007. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. Journal of Derivatives 15, 52–66.

Boogert, A., de Jong, C., 2011. Gas storage valuation using a multifactor price process. Journal of Energy Markets 4, 29–52.

Cartea, A., Williams, T., 2008. Uk gas markets: The market price of risk and applications to multiple interruptible supply contracts. Energy Economics 30, 829–846.

Chen, Z., Forsyth, P.A., 2006. Stochastic models of natural gas prices and applications to natural gas storage valuation. Working paper.

Cortazar, G., Schwartz, E.S., 2003. Implementing a stochastic model for oil futures prices. Energy Economics 25, 215–238.

Jaillet, P., Ronn, E.I., Tompaidis, S., 2004. Valuation of commodity-based swing options. Management Science 50, 909–921.

Schwartz, E.S., 1997. The stochastic behaviour of commodity prices: Implications for valuation and hedging. The journal of finance 52, 923–973.

Schwartz, E.S., Smith, J.E., 2000. Short-term variations and long-term dynamics in commodity prices. Management science 46, 893–911.

Stoll, S.O., Wiebauer, K., 2010. A spot price model for natural gas considering temperature as an exogenous factor and applications. Journal of Energy Markets 3, 113–128.

Weron, R., 2006. Modeling and forecasting electricity loads and prices: A statistical approach. John Wiley & Sons Ltd, Chichester.

Xu, Z., 2004. Stochastic models for gas prices. Master's thesis. University of Calgary.