Estimating the value of flexibility from real options: On the accuracy of hybrid electricity price models

HEMF Working Paper No. 04/2018

by

Christian Pape

and

Oliver Woll

and

Christoph Weber

March 2018

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

March 2018

Estimating the value of flexibility from real options: On the accuracy of hybrid electricity price models

Christian Pape · Oliver Woll · Christoph Weber

Abstract Practitioners in the electricity industry aim to assess the value of power plants or other real options several months or even years ahead of operation. Such a valuation is notably required for hedging purposes. The revenue streams to be earned in the spot market are thereby already secured on future markets. Yet the peculiarities of the electricity market, notably the limited storability of electricity and the incompleteness of the derivative markets, make this problem also theoretically challenging since they prevent the straightforward application of standard approaches for price modeling and for hedging.

In this context, the contribution of this article is twofold: (1) We present a novel methodology to model electricity prices based on fundamental expectations and accounting for both short-term and long-term uncertainties. This requires the joint modeling of different commodity prices, namely electricity, fuel and CO_2 prices. Moreover price distributions have to be modelled in order to assess the real option value adequately ex ante. Specifically, we compare two different modeling approaches to account for long-term variations in multi-commodity price dynamics. (2) We suggest a test procedure and introduce performance measures to analyze the accuracy of the proposed price modeling. We thereby focus on the practically relevant question, whether the price modeling provides ex ante estimates of the value of the real option that are in line with the ex post realized values. This approach is chosen since no derivative markets exist where the (extrinsic) values for the real options could be observed months or years ahead of actual operation. Nonetheless we show that under well-defined assumptions, the ex-ante values derived using the price model should provide unbiased estimates of the ex post values, which are computed as a sum of hedging and spot exercise revenues.

C. Pape (✉) · C. Weber
House of Energy Markets and Finance, University of Duisburg-Essen, Germany
Universitätsstr. 12, 45117 Essen
e-mail: Christian.Pape@uni-due.de

O. Woll
Centre of European Economic Research (ZEW)
L 7, 1, 68161 Mannheim, Germany
woll@zew.de

The application part shows results for a state-of-the-art gas power plant. By applying the developed performance measures and test statistics, we find that neither of the two investigated price models clearly outperforms the other.

Keywords Electricity price forecasting · Futures market · Hedging · Real option · Stochastic optimization· Valuation

1 Introduction

In electricity markets, the valuation of flexibilities from conventional generation assets and the adequate description of short- and long-term electricity price uncertainties are of substantial interest for market participants to hedge against the related price and volume risks. The expected revenues in the spot market are thereby secured on future markets several months or even years ahead of actual operation. In recent years the valuation and hedging became increasingly challenging due to changes in the industry such as the increasing renewable energy production or decreasing electricity price levels. We subsequently first discuss the challenges in terms of electricity price modelling and then those in terms of computation and validation of real option values. Those two issues motivate our approach for electricity price modelling and the evaluation methodology developed to assess the performance and unbiasedness of real option values.

Whereas longer-term electricity price models estimate the price development for time horizons of one to several years (e.g., to schedule power plant revisions), shorter-term electricity price models aim to forecast for time horizons ranging from several hours to less than a week (e.g., to optimize the day-ahead unit commitment). Due to the increasing share of fluctuating renewable infeed, the consideration of short-term stochasticity in longer-term price forecasting models is crucial for the valuation of asset flexibilities, the assessment of optionalities and other risk management purposes (Wozabal et al. 2014). Asset flexibilities notably arise for power plants which can be switched on and off under some technical constraints, i.e. these are real options exercised against spot electricity prices. The value of these real options then depends on the spread between spot electricity prices and variable generation cost of the power plant. For conventional thermal generation units, the variable generation cost in turn depend on input factor prices, namely fuel (e.g. gas) and CO_2 certificate prices. Hence a power plant may be viewed as a real option with multiple underlying commodities.

Yet, the long-term relations among stock or commodity prices tend to be unstable for financial markets and for energy commodities the fluctuations are even more extreme (among others Alexander 1999, Eydeland and Wolyniec 2003, Bencivenga et al. 2010).[2] Therefore, an accurate representation of the dependencies between input commodity prices and electricity prices is of importance for the evaluation of such spread options (Aïd et al. 2013). In this article, we compare two electricity price models to consider the long-term uncertainty. Model 1 uses a one-factor mean reversion process to directly model the longer-term clean generation spread between electricity price and the combined price for the input fuel and CO_2 emissions,[3] whereas Model 2 adopts a multi-factor approach to describe the longer-term variations of electricity prices and input commodity prices.

Usually, the accuracy of electricity price forecasts is assessed by performing statistical tests to measure the accuracy of point, interval and distributional forecasts (among others see Janczura and Weron 2012; Weron 2014). Statistical tests typically compare the price forecasts with in-sample or out-of-sample realizations of prices. However, the price forecasting accuracy does not directly measure the adequacy of estimates for the real option value. A model that is judged superior in terms of price forecast accuracy may well underperform when options on price spreads with multiple underlying factors are to be assessed. At the same time, the ex-ante value of such real options is generally not observable in the market. Based on readily available futures market quotes, only the current average generation spread for low time granularities (e.g. months) may be computed. This yet provides just a rough estimate of the intrinsic option value given the combined effects of market incompleteness and limited storability. Limited storability induces strong price fluctuations in the spot market e.g. between day and night hours or between periods with low and high renewable infeed. Due to the market incompleteness of the futures markets which quote average prices for longer delivery periods (e.g. months), the short-term price fluctuations are yet averaged out in the futures markets. Correspondingly the real option value of switching on and off is not directly revealed in the derivative markets with lower time granularity.

Given this market incompleteness, it is also not straightforward to construct a dynamic delta hedge and such a hedge will moreover not fully eliminate the price risk. Under well-defined assumptions discussed below, ex ante option values obtained from price simulations will nevertheless be an unbiased estimate of the ex post realized value. The latter may be computed as the sum of the cash flows from hedging operations plus cash flows from an optimal dispatch in the spot market. Yet an in-depth discussion is necessary to understand how differences between ex post and ex ante values are related to market incompleteness and limited storability on the one side, and inaccuracies of the price models on the other.

---

[2] Among others, Guo et al. (2014) highlight the importance of correlation risk for spread options, thereby referring to the hedging of a gas-fired power plant. Related to more general spread hedging, Carmona and Durrleman (2003) provide notable work on multi-asset hedging.

[3] The term ‘clean spark spread’ designates a spark spread (the spread between electricity and gas prices) that includes emission costs. We use the general term ‘clean generation spread’ because the Models presented in this paper are also applicable for other spreads, e.g. ‘clean dark spread’ (the spread between electricity and coal prices and emission costs).

The aim of this article is hence to assess the value of real options in the electricity market and to what extent an inaccurate pricing model could cause valuation biases. To understand the accuracy of said pricing models, we suggest an evaluation framework that focusses on the accuracy of longer-term distributional forecasts and attempts to decompose the impact of model imperfections. The interplay with market imperfections is thereby scrutinized in detail.

The contribution of this article is twofold: (1) We present a novel methodology to model electricity prices based on fundamental expectations and accounting for both short-term and long-term uncertainties. As the literature on joint modeling of commodity prices is scarce, we compare two modeling approaches to account for variations in long-term price dynamics (Kovacevic and Paraschiv 2014). (2) We suggest a test procedure and introduce performance measures to analyze the accuracy of the proposed price modeling for estimating the value of flexibility from real options. The evaluation framework takes into account that prior to delivery only the futures market can be used for hedging.[4]

The article is structured as follow: **Section 2** introduces the two electricity price models under study. The section describes the fundamental price modeling, the modeling for the short-term uncertainty and the differences between the two methods to capture the long-term stochastics of electricity prices. **Section 3** describes the valuation model used to calculate the real option value ex ante. **Section 4** introduces the evaluation framework to assess the accuracy of electricity price forecasts for valuation purposes. In this section, we explain the hedging with futures and propose novel performance and test measures. **Section 5** presents a real-world application for a large-scale gas power plant. For both price models we perform the asset valuation and apply the suggested evaluation framework. **Section 6** concludes with a summary of our main findings.

2 Hybrid energy price modeling

In addition to the forecast horizon, electricity price models are classified based on the method applied (Weber 2005; Weron 2014). We introduce a so-called hybrid model to combine the strengths of different modelling approaches, namely, a fundamental and a stochastic method. Fundamental models attempt to capture physical and economic relationships in electricity markets and are beneficial for depicting longer-term changes in supply and demand. Such models are parameter-

[4] Additionally, this article has practical relevance. Typically, utilities use sophisticated models to derive expectations regarding longer-term electricity prices based on a few number of scenarios. Normally companies use one leading scenario which is in line with all strategic decisions and two competing scenarios which reflect an up- and a downside case. These expectations form the basis for corporate planning and are approved by top management. The price modeling presented in this paper provides a methodology to consider uncertainty in such scenarios.

rich and based on specific assumptions regarding economic or physical relationships. They are, however, reacting sensitively if these assumptions are violated in reality (Weron 2014). Therefore, we add a stochastic component, which is advantageous for modeling the price variation. The stochastic component distinguishes between short- and long-term dynamics as they are driven by different sources of uncertainty. Thereby, the fundamental modelling approach ensures consistency between input commodity and electricity prices.[5] However, the inclusion of the longer-term spread uncertainty within the fundamental model affects the bandwidth of the forecasted price distribution.

2.1 Model definition

The hybrid price modeling is based on the assumption that electricity spot prices S_t at time t (i.e. hours) are primarily driven by fundamental information and are also subject to stochastic influences. To compare the differences in modeling the longer-term uncertainty, we define two models: Model 1 estimates a deterministic path for the fundamentally expected price $\bar{S}_t^{fund}$ based on expected input commodity prices $\bar{X}_{c,t}$ and models a short-term price dynamic ω_t^{spot} as well as a long-term spread uncertainty ω_t^{spread}, see Eq. (2-1). The index n belongs to the set of simulations $n \in \{1, \dots N\}$, with N indicating the number of simulations and the index $c \in (1,\dots,C)$ indicating different commodities.

$$S_{n,t}^{Model\ 1} = \bar{S}_t^{fund}\left(\bar{X}_{c,t}\right) + \omega_{n,t}^{spot} + \omega_{n,t}^{spread} \qquad (2\text{-}1)$$

Model 2 applies a multi-factor approach to model the input prices $X_{c,n,t}$. The resulting input prices are used to derive N estimates for $S_{n,t}^{fund}$ such that the long-term spread uncertainty is included in S_t^{fund} and ω_t^{spread} becomes superfluous.

$$S_{n,t}^{Model\ 2} = S_{n,t}^{fund}\left(X_{c,n,t}\right) + \omega_{n,t}^{spot} \qquad (2\text{-}2)$$

To keep Model 1 and 2 comparable, the short-term stochastic $\omega_{n,t}^{spot}$ and the approach to model S_t^{fund} are equal for both models such that the probability weighted means of $S_{n,t}^{Model\ 1}$ and $S_{n,t}^{Model\ 2}$ over all N are equal.[6] $\bar{S}_t$ is the notation for the mean of the simulated prices given by Eq. (2-3)

[5] Hereby, consistent means that commodity price developments are in line with fundamental relationships, e.g. increasing gas prices tend to increase the electricity prices due to higher variable costs of gas-fired power plants whereas emission prices increase the variable costs for all conventional assets based on their respective emission intensity.

[6] This assumption is true as long as variations in input commodity prices do not lead to a changed ordering of the bid stack. As discussed in Sunderkötter and Weber (2012), such reversals in the

$$\bar{S}_t = \sum_n pr_{n,t} \cdot S_{n,t}^{Model\ 1} = \sum_n pr_{n,t} \cdot S_{n,t}^{Model\ 2}, \qquad (2\text{-}3)$$

with the probability of occurrence $pr_{n,t}$ for state n. Given the homogeneous probabilities of $pr_{n,t}$ over n, Eq. (2-3) corresponds to the average of all prices $S_{n,t}$. Any time the notation S_t is used without further specification of the model, either $S_{n,t}^{Model\ 2}$ or $S_{n,t}^{Model\ 1}$ can be applied. [7] Fig. 1 summarizes and illustrates the model differences.

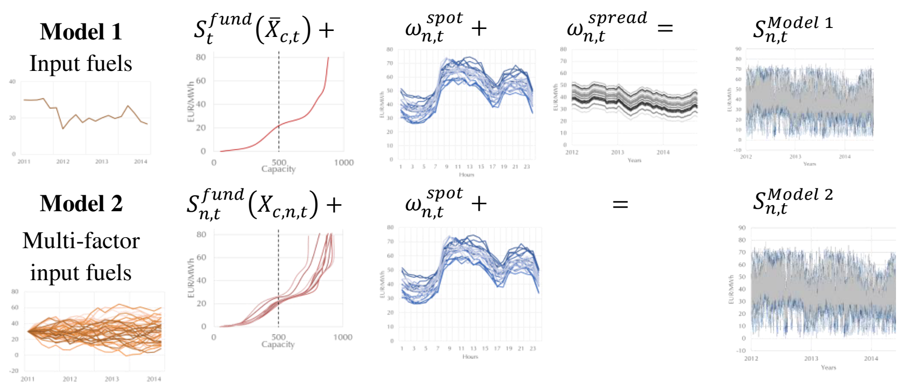

Fig. 1 Overview of the hybrid price simulation models

2.2 Fundamental model

The fundamental price $S_{n,t}^{fund}$ is determined in an equilibrium model for supply and demand for N simulations of commodity prices, similar to Müsgens 2006; Weigt and Hirschhausen 2008; Graf and Wozabal 2013; Pape et al. 2016. First, we assume that companies bid their available capacity at their operational costs $c_{u,t}$ given by

merit order are unlikely in the German market. Weber and Vogel (2014) provide evidence that even carbon constraints do not lead to reversals in the stack in long-term equilibria.

[7] In order to estimate the long-term uncertainty (Model 1: $\omega_{n,t}^{spread}$ and Model 2: $dX_{t,c}$, see section 2.4) and the short-term uncertainty, a Kalman filtering approach or similar techniques may be used. Through the use of latent variables, those approaches separate the short-term and long-term components of uncertainty. Since the short-term impact of long-term uncertainties is low, neglecting the latter in the short-term estimation is expected to induce little bias. The other way round, a simple moving average procedure on spot prices serves as a first approximate filtering approach when estimating long-term stochasticity.

$$c_{u,n,t} = \frac{\left(x_{f,n,t}+x_{CO_2,n,t}\cdot CO_{2\,f,u}\right)}{\eta_u} + c_{u,other}. \tag{2-4}$$

The operational costs are driven by time-varying fuel and emission prices $X_{f,t}$. The index $f \in \{1, \ldots, M\}$ is a subset of index c and contains the type of fuel that is used as input for any production unit $u=\{1,\ldots,U\}$. M is the number of fuel types. The notation for the production units efficiencies is η_u and $CO_{2_{f,u}}$ for emission intensities. Second, we assume that the market operator aggregates the bids and arranges them in increasing order of the operational costs $c_{u,t}(c_{1,t}<c_{2,t}<\ldots<c_{U,t})$. Given those ordered capacities, the stepwise bidding quantities $B_{u,t}$ are given by

$$B_{u,t} := \left[\textstyle\sum_{i=1}^{u-1} K_{i,t}, \sum_{i=1}^{u} K_{i,t}\right), 1 \le u \le U, \tag{2-5}$$

with the convention $\sum_{i=1}^{0} = 0$ (cf. Aïd et al. 2009). $K_{i,t} \in R_+$ represents the available capacity, thus the installed capacity minus unavailable capacity (e.g., due to power plant revisions) at time t. The operational costs sorted in ascending order will deliver the typical monotonous shape of the bid stack. The fundamental estimation of the spot price $S_{t,n}^{fund}$ at time t is given by the operational costs of the last unit that is needed to satisfy a given demand:

$$S_{n,t}^{fund} := \textstyle\sum_{i=1}^{U} cost_{i,n,t}\, \mathbb{1}_{\{D_t \in B_{i,t}\}}, for\ t \in \mathbb{N}. \tag{2-6}$$

The demand D_t denotes the residual demand. The variable renewable energy sources RES_t, e.g., wind and solar, are assumed to produce at zero variable costs and are subtracted from domestic demand L_t.

$$D_t = L_t - RES_t \tag{2-7}$$

Due to the time-varying nature of the residual load and the non-linearity of the bid stack, the basic fundamental drivers of electricity price formation are covered in the fundamental model; however, other factors such as strategic bidding are not included.

2.3 Short-term stochastic variation

Electricity prices are a result of idiosyncratic influences causing seasonal cycles, high and non-constant volatility (heteroscedasticity) and mean reversion. For the short-term variation of electricity spot prices, we model a stochastic process for the difference ω_t^{spot} between observed prices S_t and fundamental expectation $\bar{S}_t^{fund}$, Eq. (2-3). We use a discrete time approach covering four steps: (A) The day-ahead bidding is done simultaneously for all 24 hours of the day and thus based on the same information set. In line with Huisman et al. (2007), we treat the time series as a panel of individual hours $h = \{1, ..., 24\} \in t$ observed with a daily frequency $d = \{1,...,D\} \in t$. (B) As ω_t^{spot} result from comparing actual prices and prices based on a stack model, their distribution is non-normal. We map their empirical distribution to a normal distribution as $T_h: \omega_{d,h} \rightarrow u_{d,h} = \Phi^{-1}\{CDF_h(\omega_{d,h})\}$. Φ^{-1} is the inverse of the standard normal distribution and CDF_h is the empirical cumulative density function. (C) The stochastic influences on the panels are complex and interrelated. To reduce the complexity, we use a principal component analysis. We identify common principal factors $f_{t,i}$ by solving the eigenvalue-eigenvector-problem for the correlation matrix from $u_{d,h}$ (among others, see Kovacevic and Wozabal (2014), Ziel (2016), Zugno et al. (2013)). (D) The factor $f_{t,i}$ is modeled based on an ARMA-GARCH specification to cover lagged effects in price level and price volatility, which are characteristic of electricity prices. The maximum likelihood method is used to estimate the parameters. For the simulation of hourly prices, steps (A) to (D) are carried out in reverse order. The short-term uncertainty $\omega_{d,h}^{spot}$ is the same for Model 1 and Model 2. A summary of the approach and the key equations are given in Appendix A. More details and the fit with data is shown in Weber (2007) and Pape et al. (2017). Pape et al. (2017) report a high R^2 of 0.81 for Model 5 (which corresponds to the specification used here) and a low MAE for an out-of-sample day-ahead forecasting at 4.09 EUR/MWh. This underlines the good performance of this approach.

2.4 Long-term stochastic variation

Without explicitly referring to the longer-term dependencies between commodity prices, Frikha and Lemaire (2013) model simple correlations between Ornstein-Uhlenbeck processes. Empirical work on dependencies among energy commodity prices is reported by Bachmeiner and Griffin 2006; Mjelde and Bessler 2009; Joets and Mignon 2011. However, a limited number of studies have jointly modeled commodity prices, including coal, gas and emissions. To the best of our knowledge, a joint model for European commodity prices was first applied by Kovacevic and Paraschiv (2014), who use principal component analysis to identify the joint factors of commodity prices.

2.4.1 Model 1: One-factor mean reversion approach

The modeling outlined so far neglects the influence of longer-term electricity price variations. We continue with Eq. (2-1) and find a specification for $\omega_{n,t}^{spread}$ that describes the longer-term uncertainty of the clean generation spread including CO_2

emissions. This is similar in vein to the approach adopted in Frikha and Lemaire (2013) although they consider more complex mean-reversion dynamics. The main drivers of for the modelled spread are capacity shortages or oversupply in relation to the actual demand level that may change unexpectedly, e.g. due to economic crises. In the longer run, capacity adjustments will drive the spread levels back to equilibrium; therefore, mean-reversion is likely to occur. The dashed line in Fig. 2 shows the clean generation spread. It provides graphical evidence and Table 14 in Appendix B gives statistical support to the mean reversion assumption for Model 1.[8]

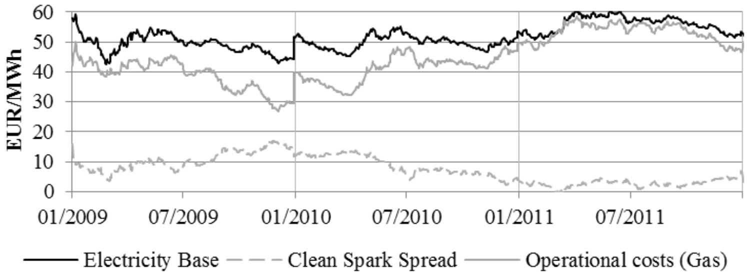

Fig. 2 Clean-Spark-Spread, operational costs and electricity prices based on front-year contracts

Above equilibrium, the spread provides incentives for capacity investments. Below equilibrium, the spread motivates capacity shut-down. The spread is in equilibrium when it enables investors to refinance the investment and fixed costs based on the expected yearly electricity production. At equilibrium, there are no incentives for market entries or exits and any divergence from the equilibrium will return to the spread equilibrium again (Woll and Weber 2011). In continuous time, the standard mean reversion process is the Ornstein-Uhlenbeck process, e.g., see Dixit and Pindyck (1994). In discrete time, the equivalent formulation is a first-order autoregressive process AR(1):

$$\omega_{n,t}^{spread} = \mu(1 - e^{-\kappa\,\Delta t}) + e^{-\kappa\,\Delta t}\omega_{n,t-1}^{spread} + \varepsilon_t. \tag{2-8}$$

The stochastic term ε_t is normally distributed with the spread volatility σ^{spread}. Parameter κ is the mean reversion rate, the speed of returning to equilibrium μ.

2.4.2 Model 2: Multi-factor approach

The modeling outlined for Model 1 is a multi-factor model by definition (cf. e.g., Gibson and Schwartz 1990; Schwartz 1997; Huisman and Mahieu 2003). However,

[8] On the consistency of mean-reversion in spot and random walk for future prices cf. Dempster et al. (2008)

with respect to the longer-term spread, it constitutes a one-factor model due to a single stochastic term in Eq. (2-8). The basic idea for the multi-factor approach is that the operational costs will cover the long-term spread dynamics if the fuel prices are modeled stochastically. A large share of fuel-price-dependent power generation supports such a hypothesis. Therefore, we treat the fuel prices $X_{n,c,t}$ as stochastic variables in Model 2. E.g., Aïd et al. (2013) pursue a similar modeling idea and model the fuel spreads as independent geometric Brownian motions GBM. Fig. 3 and Table 15 in Appendix B support the validity of GBM assumption for Model 2.

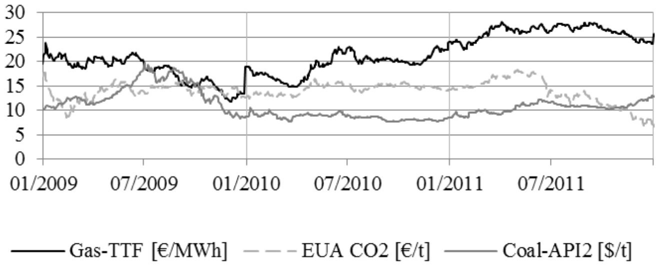

Fig. 3 Gas-TTF, EUA CO2 and Coal-API2 calendar front-year prices from 2009 until 2011

In contrast to Aïd et al. (2013) we directly model commodity prices in discrete time and account for their joint behavior in the error term. In discrete time, the equivalent formulation of a GMB is given by Eq. (2-9).

$$X_{t,c} = X_{t-1,c} e^{\left(r_c - \frac{1}{2}\sigma_c^2\right)\Delta t + \sigma_c \sqrt{\Delta t}\varepsilon_{t,c}} \; with \; \varepsilon_{t,c} \sim N(0,1). \quad (2\text{-}9)$$

Hereby, σ_c is the volatility for commodity price c, r_c the drift parameter and dW_c the increment of a Brownian motion.[9] To model the dependencies between commodity prices, we enforce a constant correlation (with respect to time) for the random variables; $\mathrm{Corr}(\varepsilon_{t,c}, \varepsilon_{t,c'}) = \rho_{c,c'}$, where $\rho_{c,c'}$ denotes the correlation of the input prices X_c and $X_{c'}$. We model $\boldsymbol{\varepsilon}_t$ as

$$\boldsymbol{\varepsilon}_t = \mathrm{L}\mathbf{z}_t, \quad (2\text{-}10)$$

where $\mathbf{z}_t$ is a vector of uncorrelated random variables and L a lower triangular matrix given by the Cholesky decomposition of a covariance matrix Σ. This matrix is positive and semi-definite by construction and a real-valued Cholesky matrix exists that satisfies $\Sigma = \mathrm{LL}^T$. $\mathrm{L}\mathbf{z}_t$ produces random variables with covariance properties of the commodity prices. Simulations for $X_{n,c,t}$ result in N bid stacks, that are used in the fundamental model Eq.(2-6).

[9] The short-term electricity price model covers effects of non-constant volatility. These tend to overshadow the effects of non-constant volatility in the input commodity prices. Therefore, and given the methodological challenges of specifying a multi-commodity non-constant volatility model, we keep the assumption of constant volatility for the input commodities and the long-term spread modelling.

3 Valuation of real options

With technical constraints, the decision to exercise a real option does not depend solely on the spot prices of electricity and input commodities but also on the plant operating state. The valuation problem becomes path dependent and complex to solve (Gardner and Zhuang 2000). The valuation model applied in this article is based on stochastic dynamic programming with backward recursion, similar to Tseng and Barz (2002) and Weber (2005). The basic model covers technical constraints for minimum operation and ramp times as well as minimum and maximum generation levels and start-up costs. The relevant formulas for the cost function, start-up costs C_t^{St} and electricity production P_t^{El} can be found in Appendix C and detailed information is provided in Gardner and Zhuang (2000), Tseng and Barz (2002) or Weber (2005).

To determine the value of flexibility, we first evaluate the intrinsic option value based on expected values for electricity prices $\bar{S}_t$ and input cost $\bar{X}_t^{total}$, i.e. the mean given by Eq. (2-3). $\bar{X}_t^{total}$ is the total cost of inputs including the prices for fuel and for emissions certificates. The intrinsic value $\bar{V}_{t,T_D}$ in t for delivery period T_D from T_1 until T_2 equals the differences of the optimal profit $\bar{V}_{t\to T_2}$ and $\bar{V}_{t\to T_1}$ at time t, with $\bar{V}_{t\to T}$ given by[10]

$$\bar{V}_{t\to T}(O'_{t-1}) = \max_{O'_t, P_t^{El}} \Big(\bar{S}_t \cdot P_t^{El} - \bar{X}_t^{total} \cdot P_t^{Fu}(O'_t, P_t^{El}) - c^{other} \cdot P_t^{El} - C_t^{St}(O'_t) + \bar{V}_{t+1\to T}(O'_t)\Big). \quad (3\text{-}1)$$

Further elements appearing in Eq.(3-1) are the consumed input fuels P_t^{Fu}, other variable costs c^{other} and start-up costs C_t^{St}. The fuel quantity and start-up costs depend on the state variable O'_t, which is a generalization of the state variable for the operation state O_t and counts the hours since start-up if positive or the hours since shutdown if negative. The revenues in Eq. (3-1) are driven by the optimal electricity output P_t^{El}, given the operation state O_t. Eq. (3-1) elucidates that $\bar{V}_{t\to T}$ depends on the value of the state variable preceding the optimization period O'_{t-1} and may be computed recursively. Therefore, the problem can be solved by finding the optimal state in the final node and working back in time. A detailed description of the procedure for backward induction is found in Weber (2005). As stated in Eq. (2-3), the intrinsic value $\bar{V}_{t\to T}$ is identical for Model 1 and Model 2 in the present study. The total (extrinsic) option value is based on the valuation with stochastic electricity price simulations $S_{n,t}$ with $n \in \{1,\ldots,N\}$. Now, the optimal profit does not depend solely on the previous operation status but also on the uncertain electricity prices. Similar to Eq. (3-1), the optimal profit $V_{t\to T}$ at time t, given the operation status O'_{t-1}, and the price S_{t-1} is given by

[10] This clarification is necessary to deal with options whose underlying covers a delivery period $[T_1 T_2]$ with $t<T_1$. Hence we define the option value in such a case as the difference between values for the two boundaries of the delivery period. E. g. for the intrinsic value $\bar{V}_{t,T_1,T_2}=\bar{V}_{t\to T_2}-\bar{V}_{t\to T_1}$.

$$V_{t\to T}(O'_{t-1}, S_{t-1}) = \frac{1}{N}\sum_{n=1}^{N}\left[\max_{O'_n, P_n^{El}}\left(S_{n,t}\cdot P_{n,t}^{El} - \bar{X}_t^{total}\cdot P_{n,t}^{Fu}\left(O'_{n,t}, P_t^{El}\right) - c_{other}\cdot P_{n,t}^{El} - C_t^{St}\left(O'_{n,t}\right) + V_{t+1\to T}\left(O'_{n,t}, S_{n,t}\right)\right)\right]. \quad (3\text{-}2)$$

The optimal operation strategy depends on the payoff in t for different price scenarios and on the expected optimal profit at further time steps given those different prices. The optimal unit commitment can again be solved through backward recursion in a dynamic programming approach. Therefore, Monte Carlo simulations for the price paths based on information available at time t = 0 until the final node T are necessary (among other Ripley 2009; Kroese et al. 2011). Instead of using all simulated paths, we generate a price lattice similar to Weber (2005) or Felix and Weber (2012) to reduce computational burden.[11] From initially N possible states for electricity prices $S_{n,t}$, we reduce the dimension based on a k-means algorithm (e.g., see Lloyd 1982). Given $k = (1, ..., K)$ price clusters and the corresponding transition probabilities $Tr_{k,t-1\to k'}$, Eq. (3-2) is written as[12]

$$V_{t\to T}\left(O'_{k,t-1}, S_{k,t-1}\right) = \sum_{k'} Tr_{k,t-1\to k'} \max_{O'_{t,k'}, P_{k'}^{El}}\left(\left[S_{k',t}\cdot P_{k',t}^{El} - \bar{X}_t^{total}\cdot P_{k',t}^{Fu}\left(O'_{k',t}, P_{k',t}\right) - c_{other}\cdot P_{k',t}^{El} - C_t^{St}\left(O'_{k',t}\right) + V_{t+1\to T}\left(O'_{k,t}, S_{k',t}\right)\right]\right). \quad (3\text{-}3)$$

The value of flexibility $V_{t\to T}^{flex}$, is the difference between the intrinsic value $\bar{V}_{t\to T}$ and the total option value $V_{t\to T}$ derived from the valuation with stochastic prices.[13]

$$V_{t\to T}^{flex} = V_{t\to T} - \bar{V}_{t\to T} \quad (3\text{-}4)$$

4 Accuracy of price models

For estimating the value of flexibility $V_{t\to T}^{flex}$, a well-performing point forecast model for electricity prices is not sufficient, as it only ensures the accurate valuation of the intrinsic option value ($\bar{V}_{t\to T}$). Rather, the choice of an electricity forecasting model should be based on a comprehensive evaluation of the ability to forecast the price distribution. Since asset valuation is a major application purpose for longer

[11] For a similar real option valuation (gas storage), Felix and Weber (2012) report small differences in the results based on least square Monte Carlo compared to recombining trees.

[12] The transition probabilities $Tr_{k,t-1\to k'}$ describe the share of simulations whereby the price in t belongs to the cluster k and in t - 1 to cluster k'.

[13] The value of flexibility as introduced here refers to the option value part that is induced by uncertainties. Basically operating flexibility may affect operations also without uncertainty and hence also the intrinsic value of a power plant may include some flexibility value, In line with the current practice in industry, we yet define the value of flexibility as the difference between the intrinsic and the total option value.

term price distribution forecasts, the accuracy of ex ante asset valuations is a practically relevant indicator for the quality of the distribution forecasts. Yet the definition and the use of such an indicator raises some challenges – notably related to the characteristics of electricity markets. Primarily, three aspects have to be considered:

(a) Since operational decisions for power plants and similar flexibilities are taken close to real time, the spot price is the reference to be used for the valuation of the corresponding real options. Yet, the electricity futures markets do not replicate the granularity of the spot markets, rather they provide quotes for delivery over longer time periods (e.g. months). Given the limited storability of electricity, the prices in spot markets vary strongly whereas these fluctuations are averaged out in the future quotes.

(b) As electricity futures are written for the delivery over a certain time period such as months, model-based estimates and simulations are needed to compute future prices in a finer granularity (cf. section 2). Therefore, we simulate hourly prices and have to average over possibly heterogeneous time segments to determine futures prices (see Eq. (4-1)). This leads to a joint testing problem: the different components of the electricity and fuel price models can only be tested jointly - here in conjunction with a valuation model, used to evaluate directly real option values.[14]

(c) Under standard theoretical assumptions, a dynamic delta hedge ensures the recovery of the initial (extrinsic) option value.[15] However, the incompleteness of markets and the impossibility of cash-and-carry arbitrage (which results from limited storability) make adaptations of the hedge strategies necessary. Only in such a framework, the ex-ante option values may be computed and compared to the ex-post values. The evaluation framework derived from these considerations is described in the next subsection. Subsequently the necessary assumptions and the impact of market imperfections are discussed. In the last subsection, performance measures and test statistics are introduced.

4.1 Evaluation framework

We propose to evaluate whether one price forecasting model is superior to another in a framework which replicates key features of power plant management in a real-world portfolio context. This approach reflects the real option

[14] Evaluating the performance of any price forecasting and valuation approach is facing the joint hypothesis problem raised in the seminal work by Fama (1970) and discussed frequently since. I.e., we cannot simultaneously test the efficient market hypothesis and a specific market price and asset valuation model. Therefore, we assume that systematic deviations between our ex-ante asset valuations and realized values are attributable to inaccurate price and/or valuation models (see section 4.2).

[15] Standard theoretical assumptions notably include that market participants have access to the same information, assets are fully divisible, trading is possible on a continuous basis until delivery and no transaction costs exist (cf. Fama (1970)).

characteristics of power plants, the incompleteness of future markets and the longer term hedging of the real option value that is realized only in the spot market. The approach is able to capture the impact of longer term uncertainties and distributional forecasts on valuation. Fig. 4 shows the evaluation framework and visualizes the used procedure. The horizontal axis is a time axis and depicts the transition from the futures to the spot market for a given delivery period. The vertical axis shows the order of steps for the evaluation. The following paragraph further explains the intuition and the steps that are summarized in Fig. 4.

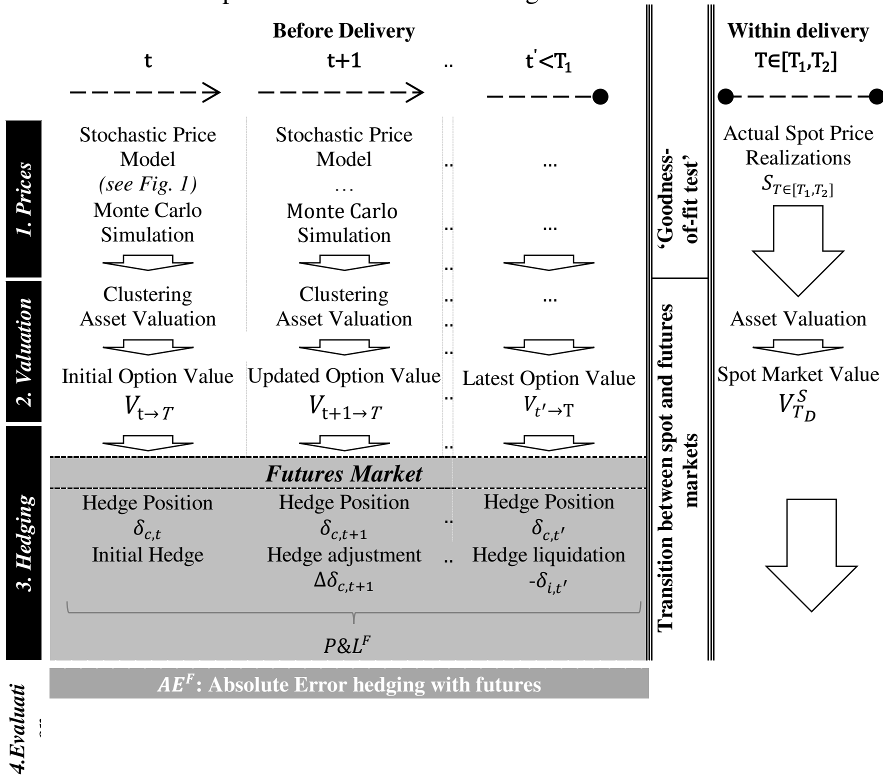

Fig. 4 Overview of the underlying problem and application of the evaluation framework.

The evaluation framework consists of two main parts. The *first part* of the evaluation framework is a straightforward application of well-developed methods to test the (statistical) 'goodness-of-fit' of electricity prices forecasts (top right of Fig. 4). Until recently, the electricity price forecasting literature neglected to focus on testing the accuracy of probabilistic forecasts (Janczura and Weron 2012; Weron 2014), but this type of research is gaining importance (among others Maciejowska and Nowotarski 2016; Pape et al. 2017). The *second part* is to define performance measures (bottom of Fig. 4) which are applicable to the hedging with futures and account for the transition between futures and spot markets. The defined performance measures are in line with standard goodness-of-fit measures and are supposed to measure the relative merits of a price model compared to another (see section 4.3). Since the focus is on the modelling of future spreads, a s pot price model is applied that fulfils Eq. (2-3) such that both models deliver the same point-forecasting results and this allows to focus on differences of the distributional forecast ability.

The starting point of the overall procedure are electricity price simulations. Based on the price simulations, we can compute the goodness-of-fit measures of part 1 (top left of Fig. 4). Before the start of delivery ($t < \mathrm{T}_1$), the simulated spot prices S_{T_D} for the delivery period of the underlying future T_D from T_1 until T_2 are calibrated to the observed electricity future quotes at time t ($F_{\mathrm{t,T_1,T_2}} = F_{\mathrm{t,T}_D}$), such that $S_{n,T}$ reflect possible spot prices in a risk neutral world, see Eq. (4-1);

$$F_{t,\mathrm{T}_D} = \frac{\sum_{n=1}^{N}\sum_{T=T_1}^{T_2} S_{n,T}}{N\cdot(T_2-T_1)}. \tag{4-1}$$

The middle of Fig. 4 shows the clustering of the price simulations and the computation of the real option values. The valuation model delivers the intrinsic value plus the value of flexibility for a delivery during the delivery period $\mathrm{T_D}$ from T_1 until T_2 based on the information at time t ($V_{\mathrm{t}\to T_D}$). The subsequently computed delta hedge should eliminate the price risk (Hull 2011). Eq. (4-2) gives the hedge position ($\delta_{\mathrm{c,t}}$) for the commodities c at time t as

$$\delta_{c,t} = \frac{\partial V_{\mathrm{t}\to T_D}}{\partial F_{c,t,T_D}}\, with\ T_D = \{T_1, \ldots, T_2\}. \tag{4-2}$$

Delta neutrality is achieved by offsetting the options delta with trades in the underlying such that the delta of the position π equals zero, e.g. at t=0 it equals

$$\pi_0 = V_{0\to T_D} - \textstyle\sum_{c\in(1,\ldots,C)} \delta_{c,0}\cdot F_{c,0,T_D}. \tag{4-3}$$

After setting up the initial hedge, the changes in the delta positions $\Delta\delta_{\mathrm{c},\tau}=\delta_{\mathrm{c},\tau} - \delta_{\mathrm{c},\tau\text{-}1}$ are traded to ensure delta neutrality at all times $\tau=\mathrm{t}+1,\ldots,\mathrm{t}'$ (dynamic hedging). Fig. 4 illustrates that we calculate the hedge position at consecutive points in time $\mathrm{t}, \mathrm{t}+1, \ldots, \mathrm{t}' < \mathrm{T}_1$ and that the hedge is liquidated before the start of delivery t'. The profit and loss from hedging ($\mathrm{P\&L^F}$) then equals

$$P\&L^F_{t,t'} = \sum_{c\in(1,\ldots,C)}\left(\delta_{c,t}\cdot F_{c,t,T_D} + \sum_{\tau=t+1}^{t'} \Delta\delta_{c,\tau}\cdot F_{c,\tau,T_D} - \delta_{c,t'}\cdot F_{c,t',T_D}\right) with\ t' < T_1. \tag{4-4}$$

If we assume continuous hedging in complete markets without transaction costs and based on the correct price model, the initial option value $V_{\mathrm{t}\to T}$ equals to the option value $V_{t'\to T}$ in t' plus the $P\&L^F$ from time t until t'.

$$V_{\mathrm{t}\to T} = P\&L^F_{t,t'} + V_{t'\to T}\, for\ t' < T_1. \tag{4-5}$$

Eq. (4-5) holds since the portfolio of the real option and the continuously adjusted dynamic hedge is risk free at any moment. Hence, its value evolves according to the risk free interest rate - which is assumed to be (close to) zero.

4.2 Underlying assumptions and market imperfections

The equality (4-5) relies on the following assumptions: (1) complete markets (i.e. one hedge product per risk factor), (2) no transaction costs, (3) continuous hedging, (4) a risk-free discount rate close to zero, (5) the use of the true price model when executing the hedge transactions. If hedging is based on an erroneous price model, the resulting portfolio will not be risk free and thus the equality is no longer guaranteed. Implicitly we assume (6) perfect or at least sufficient liquidity, i.e. no price impact of hedging transactions. If the previous assumptions hold, we may note that one assumption is not necessary: (7) absence of risk premia. Even in the presence of risk premia in the real-world measure, entering a hedge contract does not change the value of a portfolio. For the resulting portfolio, to continuously remain risk free, changes in the value of the real option (or any other derivative) will come along with compensating cash flows from hedging until the end of the hedging period, where the hedge can be liquidated without value change for the overall portfolio. If assumptions (1) to (4) and (6) hold, the difference between the left-hand side and the right-hand side of Eq. (4-5) provides an indication about the accuracy of the used price model. If this difference is observed on a sample of sufficient size, an error measure may be derived. This is indicated in the bottom of Fig. 4 and the performance measures are discussed in section 4.3.[16]

Before proceeding, we check the aforementioned assumptions against the specifics of the electricity market. For electricity markets, the assumption (1) regarding market completeness is violated due to the low (e.g. monthly) granularity of products in the futures markets compared to spot trading. Assumption (3) on continuous hedging is certainly not fulfilled in the interval between the last trade in the future market (e.g. end of the preceding month) and the actual delivery (e.g. middle of the current month). These two specificities imply that we cannot rely on the standard result for complete markets that derivative values obtained under the risk neutral measure are also valid under the real world measure. Instead of deducing this result from the basic model assumptions, we therefore state it as an additional assumption (8): Valuation results for derivatives obtained under the risk neutral measure are still valid under the real world measure. This assumption is retained instead of assumption (7) on absence of risk premia but it cannot be fully tested.[17]

Finally, Assumption (2) on absence of transaction costs does not hold. Due to the fact that every trade leads to transaction costs, continuous adjustment of the hedge position may not be optimal in real world markets (Eydeland and Wolyniec

[16] Note that these performance measures assess the accuracy of the price models in view of a specific task, namely the longer-term valuation of real options. This makes the difference compared to the goodness-of-fit measures in part one of the evaluation methodology (cf. above).

[17] Alternatively, one may maintain the more standard assumption of complete markets for the theoretical analysis. In that case any application has to discuss against prima facie evidence (hourly vs. monthly products), why obvious market incompleteness does not invalidate results

.

2003). It is beyond the scope of the present paper to identify optimal hedging strategies in the presence of transaction costs. In the view of real-world applicability it is however important to assess the hedging performance including transaction costs. Transaction costs are therefore modeled as bid-ask-spread assuming that futures can only be bought at ask F^{Ask}_{c,t,T_D} and sold at bid prices F^{Bid}_{c,t,T_D} (e.g., Gondzio et al. 2003). Given a mean price F_{c,τ,T_D} as $F_{c,t,T_D}=(F^{Ask}_{c,t,T_D}+F^{Bid}_{c,t,T_D})/2$, with $F^{Ask}_{c,t,T_D} \geq F^{Bid}_{c,t,T_D}$, the transaction costs over the hedging period turn out to be:

$$TC = \sum_{i\in c}\Big[\delta_{i,t}\mathbb{1}_{\{\delta_{i,t}>0\}}\cdot(F_{i,t,T_D}-F^{Bid}_{i,t,T_D}) + \sum_{\tau=t+1}^{t'}\Delta\delta_{i,\tau}\mathbb{1}_{\{\Delta\delta_{i,\tau}>0\}}\cdot (F_{i,\tau,T_D}-F^{Bid}_{i,\tau,T_D}) + \delta_{i,t'}\mathbb{1}_{\{\delta_{i,t'}>0\}}\cdot(F_{i,t',T_D}-F^{Bid}_{i,t,T_D}) + \delta_{i,t}\mathbb{1}_{\{\delta_{i,t}<0\}}\cdot (F^{Ask}_{i,t,T_D}-F_{i,t,T_D}-F^{Bid}_{i,t,T_D}) + \sum_{\tau=t+1}^{t'}\Delta\delta_{i,\tau}\mathbb{1}_{\{\Delta\delta_{i,\tau}<0\}}\cdot(F^{Ask}_{i,\tau,T_D}-F_{i,\tau,T_D}) + \delta_{i,t'}\mathbb{1}_{\{\delta_{i,t'}<0\}}\cdot(F^{Ask}_{i,t',T_D}-F_{i,t',T_D})\Big]. \tag{4-6}$$

4.3 Performance measures and test statistics

A simple rearrangement of Eq. (4-5) leads to a first performance measure, the absolute error for the hedging with futures AE^F as given in Eq. (4-7).

$$AE^F = \left|P\&L^F_{t,t'} + V_{t'\to T_D} - V_{t\to T_D}\right| \; for\; t' < T_1, T_D = \{T_1, \dots, T_2\} \tag{4-7}$$

Smaller values in this performance measure indicate that the price model and the corresponding valuation are more accurate. But this observation is obviously subject to some stochastic noise. By making use of Eq. (2-3) in our application, we can attribute observed differences in this performance measure primarily to the spread modelling in the long-term part of the Model 1 and Model 2.[18]

As formulated, AE^F is only applicable during a time span from the initial valuation until any time prior to the start of delivery. This implies that the final value of the real option is derived based on the price model. To avoid such a model-based accuracy measure, we derive the terminal value based on the cash flows from exercising the option, i.e., running the power plant. This leads to the absolute error spot AE^S as a straightforward extension of AE^F into the spot period:

$$AE^S = \left|P\&L^F_{t,T_1-1} + V^S_{T_D} - V_{t\to T_D}\right| \; with\; T_D = \{T_1, \dots, T_2\} \tag{4-8}$$

The exercise value in the spot market $V^S_{T_D}$ is computed by solving the valuation model based on the actual spot price realizations S_T.[19] Since assumptions (1) and

[18] Yet given that other assumptions are also violated in the application, there might be some impacts overshadowing (stochastically) the influence of the long-term price modelling.

[19] As $n = 1$ for S_T, Eq. (3-1) and (3-2) deliver the same result. Note that within the delivery period no adjustments of the hedge position are possible as the futures market closes prior to T_1.

(3) are not fulfilled, we consider the case where assumptions (2), (4) to (6) and (8) are valid. Then AE^S would be zero if the price model enabled perfect forecasts at the end of the future trading. Given remaining uncertainty at the end of the future trading period, AE^S will always be positive in our model setting. Notably, AE^S will include differences in revenues that result from realizations of risk factors that are not hedgeable. If assumption (8) holds, these differences have zero mean and will hence not bias the result. We cannot fully test that assumption, but may assess the validity of the assumption based on the following considerations: risk premia related to non-priced (and thus non-observable and non-hedgeable) risk factors are likely to induce some systematic difference between the observed future price at $t' = T_1 - 1$ and the actual spot market prices. I.e. if we do not find evidence that the average difference $F_{t',T_D} - \sum_{T_D=T_1}^{T_2} S_{T_D} /(T_2\text{-}T_1)$ is different from zero, assumption (8) is likely to hold.[20]

To further investigate the effects resulting from the discontinuity at the end of the future markets, we introduce a third performance measure which compares the cash flows from the spot market to the latest valuation result. The valuation at time T_1-1 is based on information just before delivery and delivers unbiased results if the futures prices are unbiased estimates for the spot. This measure is labeled AE^V - absolute error valuation.[21]

$$AE^V = \left|V_{T_D}^S - V_{T_1\text{-}1\to T_D}\right|. \tag{4-9}$$

In addition to the absolute measures, we calculate a performance ratio (PR) for the hedging and exercising activities.

$$PR = \frac{P\&L_{t,T_1-1}^F + V_{T_D}^S}{V_{t\mapsto T_D}}. \tag{4-10}$$

The numerator of PR is the right side of Eq. (4-8) except for $V_{t\mapsto T_D}$ which is used as the denominator. A PR equal to one indicates that the ex-ante value is perfectly recovered ex-post. Any deviation from one may be attributable to market imperfections or inadequate modeling. Transaction costs are subtracted from the $P\&L^F$, and thus from the right hand of Eq. (4-4). The index TC indicates the consideration of transaction costs in the performance measures (e.g. AE_{TC}^F, AE_{TC}^S or PR).

The incompleteness of the electricity market is associated with insufficient granularity of available hedge products in the futures market, which is another driver for imperfect hedging. For the performance measures, every traded product will deliver one sampling point (i.e. calendar, quarter or monthly products). Hence the sampling error will be lower when investigating the hedge product with the

[20] A t-test on the monthly price data used in the application does not reject the hypothesis of zero mean for the difference. And thus the assumption (8) that valuation results obtained under the risk free measure are also valid under the real world measure has some plausibility.

[21] For $t' = T_1 - 1$, $AE^S \leq AE^F + AE^V$, due to the triangular inequality |a+b|≤|a|+|b|.

highest possible granularity and evaluating the mean over all AE for different delivery time periods $T_D = \{T_1,...,T_2\}$.[22] We label the resulting values as MAE^F, MAE^S and MAE^V, respectively. The aggregated measures allow additional conclusions. Notably market incompleteness by itself can be one of the main drivers for higher values of AE^V, due to non-anticipated and non-hedgeable new information for individual spot hours. The corresponding mean error MAE^V should however tend to zero, if the short-term electricity price model is adequate. Inadequate long-term price models should rather translate into absolute errors in the future market MAE^F if the time interval between final future trading day and end of the delivery period is sufficiently short and if MAE^V is small. In addition to the MAEs, we suggest to use a two-tailed t-test for identifying statistical significant biases in the valuation outcomes. Instead of computing the absolute value in Eq. (4-7)-(4-9), we may directly use the corresponding errors labelled as $E^F_{T_D}$, $E^S_{T_D}$, $E^V_{T_D}$ and test if the errors systematically differ from zero. A rejection of the null hypothesis would imply that the corresponding price model induces biased results in the valuation.

5 Application

5.1 Data

The application evaluates the hedging and valuation of a gas-fired power plant for the calendar year and monthly products for 2012 and 2013 (see Table 1). The technical parameters for the power plant are given in Table 2. The constraints for minimum off time and the ramp time are equal to the smallest Δt used for the optimization, such that the constraints are not binding. The relevant time constraint is for minimum operation time. Other important technical parameters are start-up costs and operation level dependent fuel consumption (see Section 3).

Table 1 Overview of applications

Delivery period	Start of hedging	Case
01.01.2012 until 31.12.2012	First Wednesday 2011 (05.01.2011)	Cal-2012
01.01.2013 until 31.12.2013	First Wednesday 2012 (03.01.2011)	Cal-2013
01.01.2012 until 31.01.2012	First Wednesday in January 2011	Jan-12
…	…	…
01.12.2012 until 31.12.2012	First Wednesday in December 2011	Dec-12
…	…	…

[22] Usually, products with low granularity are liquidly traded far away from maturity which lowers the impacts from potential limited liquidity and vice versa for high granularity products.

Table 2 Gas power plant parameters

Nominal capacity	Minimum output	Efficiency at nominal capacity	Efficiency at minimum output	Minimum operating time	Start-up costs
800 MW	200 MW	0.58	0.5	6 hour	10 kEUR

All data are collected from commercial data sources; an overview is given here:

Table 3 Input data and method for describing expectations

Data[23]	Source	Product	Expectations
Load	entsoe.eu	Hourly load values (adj. to monthly consumption)	Simulation of D_T similar to Woll and Weber (2011)
RES (Wind and Solar)	eex-transparency.com	Ex-ante power production	
Plant capacities	EWL database	(based on Platts, etc.)	-
Electricity price	energate.de	Day-Ahead Auction	Simulation Model 1 vs. Model 2
Coal price	"	API#2 (CIF ARA)	
CO_2 price	"	EU CO2 Allowances	
Gas price	"	Gas-TTF	

For the hedging, products with different granularity are available in the futures market. The standard products for electricity are monthly, quarterly and yearly products for base contracts with different maturities (e.g., yearly products for 8760 hours). The futures market products for gas are monthly, quarterly, seasonal (summer/winter) and yearly products. For coal, the futures market products include monthly, quarterly and yearly products. CO_2 emission allowances are only traded as yearly products. Usually, products far from maturity show lower liquidity compared to those close to maturity. Moreover, the products with higher granularity are only traded several months or quarters before delivery. Therefore, we define a basic rule for the transition among different hedging products. Half a year before maturity, one shifts hedge quantities from yearly to quarterly products, and three months before delivery, hedge quantities are swapped from quarterly to monthly products for electricity products. Table 4 illustrates the products used for the hedging of the delivery year 2012 for 01.01.2011, 01.06.2011 and 01.03.2012. For the application with monthly delivery periods, the initial hedge is performed 52 weeks before the start of delivery. Missing price notations for monthly products are derived from the already traded products with lower granularity.

Due to data limitations, we cannot observe the bid-ask spread for all relevant products. Therefore, we extract the average historic bid-ask-spread and account for decreasing spreads for products closer to maturity. For the dynamic hedge, we assume a weekly rebalancing of the hedge position. We approximate the delta by

[23] Other than the listed fuel prices for hard coal, uranium and oil are based on assumptions because prices for hard coal depend on local extraction costs, prices for uranium are not quoted.

changing the underlying price $F_{c,t,T}$ by a small amount ($\Delta F_{c,t,T}$ = 0.1 EUR) up and down and determine the values $V_{t,T}^{up}$ and $V_{t,T}^{down}$, see Eq.(5-1) .

$$\delta_{c,t} = \frac{\partial V_{t,T}}{\partial F_{c,t,T}} \approx \frac{V_{t,T}^{up} - V_{t,T}^{down}}{2 \cdot \Delta F^{i}} \tag{5-1}$$

Table 4 Transition between different hedging products for delivery: Cal-2012

	01.01.2011			*01.06.2011*			*01.03.2012*		
Products	Year	Quarter	Months	Year	Quarter	Months	Year	Quarter	Months
Electricity Base	Cal12	-	-	-	Q1/12 to Q4/12	-	-	Q3/12 Q4/12	Apr12; to Jun12
Gas	Cal12	-	-	-	Q1/12 to Q4/12	-	-	Q3/12 Q4/12	Apr12; to Jun12
Coal	Cal12	-	-	-	Q1/12 to Q4/12	-	-	Q3/12 Q4/12	Apr12; to Jun12
CO2	Cal12	-	-	Cal12	-	-	Cal12	-	-

5.2 Price simulations

Table 5 shows the estimation results. The parameters for Model 1 are estimated using the least squares regression method (Hamilton 1994). For Model 2, we calculate the covariance matrix and perform the Cholesky decomposition.

Table 5 Drift, volatility and mean reversion rate of daily spread (Model 1) and drift, volatility and correlation coefficient of weekly log-price changes (Model 2) based on data for 2009-2011 (Units of original time series Spread in €/MWh, Coal in €/tSKE, Gas in €/MWh and CO2 in €/t).

Parameter Model 1	Spark-Spread	Parameter Model 2	Coal	Gas	CO2
μ	0	r_c	0	0	0
κ	0.0065	σ_c	0.026	0.034	0.030
σ^{spread}	0.4754	$\rho_{coal,\cdot}$	1	0.666	0.252
		$\rho_{gas,\cdot}$	0.666	1	0.359
		$\rho_{CO_2,\cdot}$	0.252	0.359	1

For the price forecasts we run Monte Carlo simulations to obtain the approximate price distribution. Table 6 and Fig. 5 show exemplary results. Circa one year ahead of the delivery period, the probability density function (pdf) of the simulated prices is wider than the pdf of realized spot prices due to the high uncertainty about electricity spot prices far ahead of delivery and due to futures market quotes which indicate that market expectations were above the realized price level. The non-linearity of the supply stack (steeper at higher price levels) further caused a wider distribution of the simulated prices. The price distribution from Model 2 shows a slightly wider bandwidth compared to Model 1, as indicated by Fig. 5 and Table 6 and a standard deviation of 27.58 EUR/MWh, compared to Model 1 with a standard deviation of 25.31 EUR/MWh. The tendency of wider price distributions for Model 2 holds for nearly all times of re-estimation. The out-of-sample goodness-of-fit indicates a better performance of Model 1 compared to Model 2 (see Appendix D, Fig. 7). The graphical test in Fig. 6 going back to Rosenblatt (1952) and also known as probability integral transform (PIT), shows uniformity of the transformed values for an ideal distribution forecasts (see Diebold et al. (1998) and).

Table 6 Quantile from the price simulation for cal. 2012 based on information at 05.01.2011.

Quantile	0%	10%	20%	30%	40%	50%
Model 1	-513.29	22.93	31.40	37.73	43.68	49.77
Model 2	-513.79	20.76	29.76	36.63	42.95	49.32
		60%	**70%**	**80%**	**90%**	**100%**
Model 1		56.52	64.42	73.62	85.26	250.39
Model 2		56.19	64.06	73.68	87.75	284.91

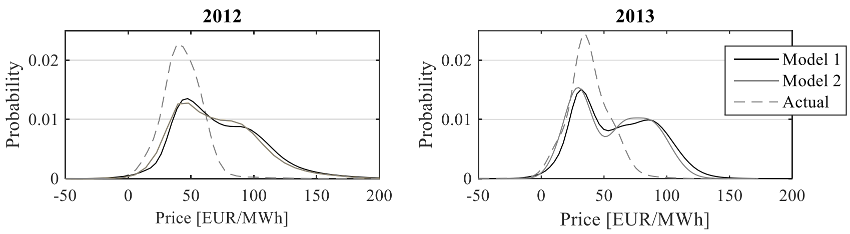

Fig. 5 Probability density function of prices for cal. 2012 and 2013 in Jan. of the previous years

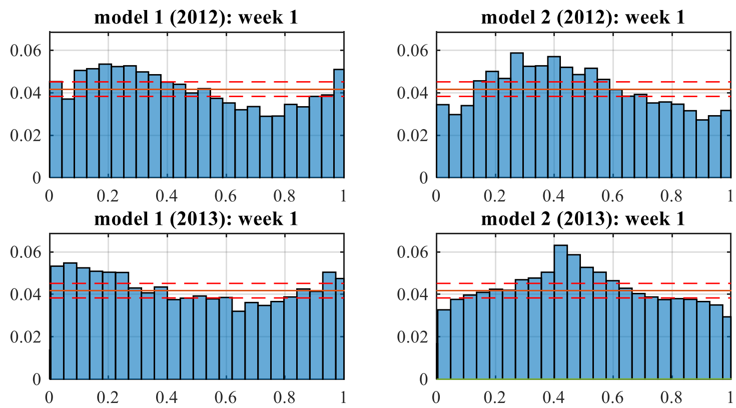

Fig. 6 Goodness-of-fit test: PIT values of price simulations from week 1 of valuation for 2012 and 2013 using Model 1 and Model 2

In addition to the graphical test, we calculated the mean CRPS (Continuous Ranked Probability Score)[24] and used the Diebold-Mariano-type test to check the statistical significance of differences in the predictive performance. Table 7 and Table 16 summarize the results by showing the number of instances at a 5% significance level in which Model 1 or Model 2 is the preferred model and how often we reject the null hypothesis which implies equal predictive performance.

Table 7 Summary of Diebold-Mariano-test of equal predictive performance for 2012 and 2013

	Preferred Model 1	**Preferred Model 2**	**Equal predictive performance (rej. H0)**
DM test (α = 5%) as share of weeks (Sample: Week 1-50, 2012)	50.0%	42.2%	7.8%
DM test (α = 5%) as share of weeks (Sample: Week 1-50, 2013)	41.0%	40.3%	18.7%

5.3 Valuation results

Table 8 and Table 9 show the valuation results at the time of the initial hedge. The results for the intrinsic valuation are identical for Model 1 and Model 2 due to Eq. (2-3). The valuation based on stochastic prices shows similar operating hours and fuel consumption for Model 1 and Model 2, but the operating margin is considerably higher for Model 2. The same holds for the delivery year 2013; however, compared to the delivery in 2012, the plant value is lower. Due to the decreasing spread, the operating hours decrease and result in 30% lower real option values. For the following 100 weeks, the plant value is reassessed with updated information. The results are shown in the Appendix F (Fig. 9).[25]

Table 8 Valuation results of the first valuation (05.01.2011) for delivery 2012

Results week 1	**Model 1**	**Model 2**

[24] Calibration and sharpness are different criteria for the accuracy of distributional forecasts. The Continuous Ranked Probability Score (CRPS) assess calibration and sharpness simultaneously. For more information, we refer to Gneiting et al. (2007) or Pape et al. (2017).

[25] Fig. 11 and Fig. 12 show the evolution of the hedge positions over the investigated time spans which illustrates similar patterns for both models and years. Larger changes in the evolution of hedge quantities occur either due to substantial changes in the respective underlying commodity price or at times where transitions between different hedge products occur (e.g., calendar to quarterly products). The evolution of the hedge position for monthly delivery periods can be found in Fig. 13 to Fig. 16. The driving factors for the monthly periods are identical to the ones from the application with calendar year products. Also, the hedge amounts are similar. However, Model 2 shows slightly smaller hedge amounts, which in turn results in different profit and loss from the hedging.

	Intrinsic	Total	Total
Value [Mio. EUR]	27.89	34.81	39.73
Flexibility value		6.92	11.85
Operating hours [h/year]	3'898	3954	3'893
Average operation margin [EUR/MWh$_{el}$]	8.96	11.28	13.13
Fuel consumption [GWh$_{th}$/year]	5'301	5'323	5'222

Table 9 Valuation results of the first valuation (04.01.2012) for delivery 2013

Results week 1		**Model 1**	**Model 2**
	Intrinsic	Total	Total
Value [Mio. EUR]	18.12	22.56	27.33
Flexibility value		4.44	9.21
Operating hours [h/year]	2'845	3'035	2'871
Average operation margin [EUR/MWh$_{el}$]	7.99	9.44	12.13
Fuel consumption [GWh$_{th}$/year]	3'913	4'124	3'888

Table 10 sums up the hedge results and performance measures. The option value in the spot market is considerably lower compared to the valuation results in the first week. In the application, the AE measures indicate a better performance of Model 1 for the delivery year 2012. None of the two models ensures the recovery of the initial plant value, but the performance ratio of Model 1 is 10 percentage points higher compared to Model 2; i.e., 77% of the initial plant value is actually recovered through hedging and spot operations (see Table 10 and Appendix D). The results for the delivery year 2013 differ slightly. For both models, the performance ratio increases beyond 100%, although a PR equal to one would be optimal. Errors from the price simulation model (section 5.2) are carried into the valuation model. As discussed in section 4 we can therefore only measure the relative merits of one price forecasting model compared to another and need to interpret the outcomes of different performance measures for the valuation in conjunction. In the case of Model 1, the plant operator gains from holding a short position in the futures market constantly over the entire period. Model 1 has a lower $\mathrm{AE}_{\mathrm{TC}}^{\mathrm{F}}$ than Model 2 but the AE^{V} and $\mathrm{AE}_{\mathrm{TC}}^{\mathrm{S}}$ for Model 2 indicate a better match between the last valuation and the spot market for this particular year. In four out of the six AE measures, the hedging results for annual contracts based on Model 1 hence outperform those based on Model 2. This is mainly due to the tendency of Model 2 to deliver very high plant values which are not recovered in the future market. Yet, the differences are small and the investigation of the mean errors (MAE_{TC}^{F}, MAE_{TC}^{S} and MAE_{TC}^{V}) for separate months is required for a more thorough assessment.

Table 10 Hedge results for Model 1 and 2 at time T in Mio. EUR for delivery 2012 and 2013

Hedging results [Mio. EUR]	**2012**	**2012**	**2013**	**2013**
Total option value	**Model 1**	**Model 2**	**Model 1**	**Model 2**
(a) First week (V_{t,T_1,T_2})	**34.80**	**39.72**	**22.56**	**27.34**
(b) Last week (V_{t',T_1,T_2})	**20.05**	**21.49**	**5.07**	**5.78**
Diff. value (b) – (a)	-14.75	-18.24	-17.50	-21.56
(c) Cash flow spot market ($V^S_{T_D}$)	**10.57**	**10.57**	**6.48**	**6.48**
Diff. to last week (c) – (b)	-9.47	-10.91	1.41	0.70
Cash flow from hedge transactions				
(i) Base	28.43	26.94	19.78	20.37
(ii) Gas	-4.31	-3.82	2.01	1.86
(iii) Emission certificates (CO2)	-7.83	-7.19	-0.49	-0.69
(iv) Transaction costs (TC)	-4.41	-4.16	-2.67	-2.45
(d) P&L^F incl. TC (sum of (i) – (iv))	**11.89**	**11.76**	**18.63**	**19.09**
Absolute error futures (AE^F_{TC})	**2.86**	**6.48**	**1.13**	**2.47**
Absolute error incl. spot (AE^S_{TC})	**12.34**	**17.39**	**2.54**	**1.77**
Absolute error valuation (AE^V)	**9.47**	**10.91**	**1.41**	**0.70**
Performance ratio (PR_{TC})	**77%**	**67%**	**123%**	**103%**

The results for the monthly delivery periods are given in Table 11 and Table 12. The plant values decrease over time (first to last week). Two effects contribute to this decrease: one is the loss in time value common to all options when they approach expiry. The other is the decline in the power price levels and average spreads during 2012 and parts of 2013. The latter effect also explains why the cash flow in the spot market is lower than the last option value for the majority of months in 2012.

In 2013, the plant values fall below seven-digit numbers and the spot cash flow is on average higher than the last option value. These observations also must be taken into account when examining the AE and MAE values. Model 2 estimates higher option values than Model 1 in all but three cases (first week 3/2012 as well as last week 1/2012 and 7/2012). The PR scores of Model 1 and Model 2 are distorted by the low (close to zero) cash flows from the spot market. The imperfect hedging strategy, together with low option values in 2013, lead to PRs that are greater than 100% - up to 259% for Model 1 in 2013. All MAEs show lower values for Model 1 compared to Model 2, indicating a better performance regarding the hedge and forecasting of the long-term uncertainty.

For the hedging including the spot market results, Model 1 delivers a substantially lower mean absolute error (MAE^S_{TC} 0.64 compared to 0.72). If one focuses on the sum of the errors instead of absolute errors, the total error is -3.32 Mio. EUR for Model 1 and -11.00 Mio. EUR for Model 2.

Table 11 Hedge results for Model 1 for all calendar months in 2012 and 2013

Unit: [Mio. €]	*Total option value*		*Spot cash flow*		*Cash flow of hedge transactions*					*Hedged value*	
Model 1	$V_{t,T}$	$V_{t',T}$	$V^S_{T_D}$	AE^V	*Base*	*Gas*	*CO2*	*TC*	$P\&L_{TC}$	AE^F_{TC}	AE^S_{TC}
01/12	6.63	7.54	0.80	**6.74**	3.74	-1.40	-1.15	-0.15	**1.05**	**1.96**	**4.78**
02/12	3.50	2.90	3.98	**1.08**	2.81	-1.33	-0.76	-0.11	**0.62**	**0.03**	**1.11**
03/12	1.76	1.69	0.40	**1.29**	2.77	-1.33	-0.75	-0.10	**0.59**	**0.51**	**0.78**
04/12	1.54	0.80	0.70	**0.10**	2.90	-0.87	-0.73	-0.07	**1.23**	**0.49**	**0.39**
05/12	1.51	0.45	0.29	**0.16**	2.20	-0.78	-0.62	-0.08	**0.73**	**0.33**	**0.49**
06/12	2.29	1.02	0.53	**0.49**	3.72	-1.33	-0.92	-0.11	**1.35**	**0.08**	**0.41**
07/12	2.19	1.47	0.22	**1.25**	3.09	-0.70	-0.47	-0.06	**1.85**	**1.14**	**0.12**
08/12	1.25	0.34	0.73	**0.40**	1.75	-0.65	-0.28	-0.05	**0.77**	**0.15**	**0.24**
09/12	1.81	0.64	0.52	**0.11**	2.18	-0.81	-0.43	-0.05	**0.89**	**0.29**	**0.40**
10/12	2.51	0.95	0.47	**0.48**	2.51	-0.74	-0.26	-0.04	**1.47**	**0.08**	**0.56**
11/12	3.62	1.42	0.67	**0.75**	2.97	-0.64	-0.18	-0.04	**2.11**	**0.08**	**0.84**
12/12	3.20	0.82	1.25	**0.43**	2.01	-0.16	0.00	-0.04	**1.81**	**0.57**	**0.14**
01/13	3.58	2.39	1.08	**1.31**	2.48	0.25	0.06	-0.09	**2.69**	**1.51**	**0.20**
02/13	2.56	1.05	0.76	**0.29**	2.56	-0.66	-0.31	-0.07	**1.52**	**0.01**	**0.28**
03/13	1.26	0.33	0.35	**0.02**	1.60	-0.39	-0.19	-0.07	**0.95**	**0.02**	**0.04**
04/13	0.55	0.07	0.50	**0.43**	0.40	0.19	0.01	-0.05	**0.55**	**0.06**	**0.50**
05/13	0.28	0.01	0.06	**0.05**	0.28	0.11	-0.03	-0.04	**0.32**	**0.05**	**0.10**
06/13	0.58	0.03	0.00	**0.03**	0.57	0.33	-0.03	-0.07	**0.80**	**0.26**	**0.23**
07/13	0.87	0.12	0.11	**0.01**	0.88	0.40	-0.14	-0.04	**1.10**	**0.35**	**0.34**
08/13	0.29	0.02	0.43	**0.40**	0.30	0.10	-0.03	-0.03	**0.34**	**0.08**	**0.48**
09/13	0.48	0.09	0.63	**0.54**	0.74	-0.13	-0.10	-0.04	**0.47**	**0.07**	**0.62**
10/13	2.31	0.43	0.62	**0.18**	1.39	-0.05	-0.17	-0.04	**1.13**	**0.74**	**0.56**
11/13	1.25	0.39	0.63	**0.24**	1.72	-0.36	-0.21	-0.03	**1.12**	**0.26**	**0.50**
12/13	0.82	0.14	1.31	**1.18**	1.14	-0.24	-0.08	-0.02	**0.81**	**0.13**	**1.30**
Sum	46.6	25.1	17.1		46.7	-11.2	-7.75	-1.49	**26.3**		
MAE_{TC}				**0.75**	1.95					**0.38**	**0.64**

The PRs (not given in the tables) are, by contrast, too high for Model 1 in 2013. The same is true for Model 2 but to a lesser extent. This result is caused by low initial option values below 1 Mio. EUR, which put more weight on relatively small differences in the hedge performance and spot stochasticity. Based on the results for calendar months, one can conclude that Model 1 performs slightly better than Model 2 with respect to all MAE measures. The biggest differences between Model 1 and Model 2 stem from the initial plant values. More conservative predictions from Model 1 for the achievable spread lead to better valuation results. The finding is particularly relevant in times of decreasing spreads. The benefits tend to disappear for plant values close to zero.

The t-test based on the errors for all investigated calendar month has limited statistical power given the sample size. The results shown in Table 13 indicate that the null hypothesis of an unbiased model is rejected at a 10% level for the error based on spot valuation $E^S_{T_D}$ in case of Model 2. For the error based on futures $E^F_{T_D}$, the null hypothesis is rejected for Model 1 at the 5 % significance level. Since $E^S_{T_D}$

measures the combined effect from hedging, valuation and operation in the spot, the results suggest that Model 1 is to be preferred. Yet, the fact that the results for $E^F_{T_D}$ are better for Model 2 suggests that there are some compensating effects between the future and spot period and that the test results are reflective of some joint modelling testing problem. Furthermore one may note that the underlying assumptions for using the t-test, notably the normality of the error distribution, are barely met. More reliable testing would hence require more general methods and an extension of the sample size – although the latter may raise challenges related to structural breaks. This is hence left for future research.

Table 12 Hedge results for Model 2 for all calendar months in 2012 and 2013

Unit: [Mio. €]	*Total option value*		*Spot cash flow*		*Cash flow of hedge transactions*					*Hedged value*	
Model 2	$\mathbf{V_{t,T}}$	$\mathbf{V_{t',T}}$	$\mathbf{V^S_{T_D}}$	$\mathbf{AE^V}$	*Base*	*Gas*	*CO2*	*TC*	$\mathbf{P\&L_{TC}}$	$\mathbf{AE^F_{TC}}$	$\mathbf{AE^S_{TC}}$
01/12	6.78	7.41	0.80	**6.61**	3.77	-1.41	-1.13	-0.16	**1.08**	**1.71**	**4.90**
02/12	3.71	2.98	3.98	**1.00**	2.81	-1.34	-0.74	-0.10	**0.63**	**0.10**	**0.90**
03/12	1.99	2.15	0.40	**1.75**	2.74	-1.28	-0.71	-0.10	**0.65**	**0.81**	**0.94**
04/12	1.49	1.06	0.70	**0.35**	2.75	-0.75	-0.67	-0.07	**1.27**	**0.83**	**0.48**
05/12	1.56	0.84	0.29	**0.54**	2.07	-0.63	-0.57	-0.07	**0.80**	**0.08**	**0.47**
06/12	2.61	1.20	0.53	**0.67**	3.32	-1.10	-0.81	-0.10	**1.32**	**0.10**	**0.77**
07/12	2.46	1.38	0.22	**1.16**	2.78	-0.56	-0.40	-0.06	**1.76**	**0.67**	**0.49**
08/12	1.41	0.42	0.73	**0.31**	1.65	-0.51	-0.26	-0.04	**0.84**	**0.15**	**0.17**
09/12	2.11	0.66	0.52	**0.14**	1.92	-0.65	-0.37	-0.04	**0.86**	**0.59**	**0.73**
10/12	3.03	1.04	0.47	**0.56**	2.31	-0.63	-0.22	-0.04	**1.42**	**0.57**	**1.14**
11/12	3.93	1.47	0.67	**0.80**	2.76	-0.54	-0.14	-0.04	**2.03**	**0.43**	**1.23**
12/12	4.01	0.90	1.25	**0.35**	1.87	-0.12	0.00	-0.03	**1.72**	**1.39**	**1.04**
01/13	3.62	2.47	1.08	**1.38**	2.43	0.25	0.06	-0.09	**2.65**	**1.50**	**0.11**
02/13	2.70	1.12	0.76	**0.36**	2.50	-0.65	-0.30	-0.08	**1.47**	**0.11**	**0.46**
03/13	1.55	0.39	0.35	**0.04**	1.58	-0.39	-0.19	-0.07	**0.93**	**0.22**	**0.27**
04/13	0.79	0.17	0.50	**0.33**	0.47	0.22	-0.01	-0.04	**0.64**	**0.02**	**0.35**
05/13	0.61	0.02	0.06	**0.04**	0.43	0.14	-0.05	-0.04	**0.47**	**0.11**	**0.08**
06/13	1.05	0.07	0.00	**0.07**	0.82	0.37	-0.05	-0.06	**1.08**	**0.10**	**0.03**
07/13	1.40	0.17	0.11	**0.06**	1.00	0.40	-0.15	-0.04	**1.22**	**0.02**	**0.08**
08/13	0.73	0.05	0.43	**0.38**	0.52	0.13	-0.05	-0.03	**0.58**	**0.11**	**0.27**
09/13	1.17	0.12	0.63	**0.51**	0.94	-0.17	-0.13	-0.03	**0.62**	**0.43**	**0.08**
10/13	2.92	0.53	0.62	**0.09**	1.31	-0.05	-0.16	-0.03	**1.06**	**1.32**	**1.24**
11/13	2.02	0.47	0.63	**0.16**	1.67	-0.36	-0.19	-0.02	**1.10**	**0.45**	**0.29**
12/13	1.42	0.20	1.31	**1.12**	1.15	-0.23	-0.08	-0.02	**0.82**	**0.40**	**0.72**
Sum	55.1	27.3	17.1		45.6	-9.86	-7.31	-1.41	**27.0**		
MAE_{TC}				**0.78**	1.90					**0.51**	**0.72**

Table 13 Two-tailed t-test for the error distribution from all calendar months in 2012 to 2013

t-test statistic	$\mathbf{E^V_{T_D}}$	$\mathbf{E^F_{T_D}}$	$\mathbf{E^S_{T_D}}$
Model 1	1.09	-2.05**	0.34
Model 2	1.41	-0.17	1.77*

Significances at the 0.05 level are labeled with ** and at 0.1 level with *.

6 Conclusion

This article highlights the importance of long-term electricity and input factor price modeling for the valuation and hedging of real (spread) options. We show that inappropriate distributional forecasting increases the risk of biased valuation results which may in turn induce risky hedging decisions and an incomplete recovery of ex-ante option values. In a longer term context, such biased price models may even lead to unwise investment decisions. Therefore, market participants and researchers should emphasize the modeling of such uncertainties. At the same time, they should further develop techniques to judge the accuracy of said price models for the valuation of asset flexibility. The latter is challenging for the electricity futures markets given the obvious market incompleteness.

In this context, we suggest a hybrid electricity price modeling approach that harnesses the benefits from different methodologies to accurately depict short- and long-term uncertainties. The price model contains three major parts: (1) A fundamental supply-stack model to forecast the fundamentally expected electricity price; (2) A stochastic approach to model the short-term fluctuations around the fundamental price, and (3) A model of the long-term fluctuations in the relevant multi-commodity spread. Thereby, a one-factor (mean-reversion) approach modelling directly the spread and a multi-factor commodity model are considered.

To evaluate the performance of the price model, we develop a stepwise evaluation procedure that accounts for the fact that the real option value is not observable before it is materialized in the spot market. The evaluation whether one price forecasting model is superior to another uses a framework which replicates key features of power plant hedging and operation in a real-world portfolio context. We use a straightforward application of well-developed methods to test the ‘goodness-of-fit’ of electricity prices forecasts. To evaluate the appropriateness of the price models in a real-world portfolio context, we additionally develop performance measures and test statistics which are applicable to the hedging with futures and account for the transition between future and spot markets. By applying the evaluation procedure, it is shown that the modeled electricity price distributions affect the total option value. Based on the comparison of the proposed performance measures, we can identify the key impacts that influence our valuation results. We conclude that modeling the mean-reversion behavior of spreads provides more conservative estimates of the flexibility value of real options than a conventional multi-commodity price model. The multi-factor approach tends to overestimate the variability of the spread between electricity and input factor prices, thus leading to an overestimation of the total option value. This overestimation then implies a reduced capability to recover the value through combined hedging and spot operation. Yet, simple statistical test indicate that the mean reversion model delivers biased results for the pure future trading period.

Some aspects of this analysis may be investigated further in future research. A promising starting point to eventually reach more general analytical results is to approximate the supply stack and drop the modeling at plant level, e.g., replacing it by a piece-wise linear function (Kallabis et al. 2016).

7 References

Aïd R, Campi L, Huu AN, Touzi N (2009) A Structural Risk-Neutral Model of Electricity Prices. International Journal of Theoretical and Applied Finance 12:925–947

Aïd R, Campi L, Langrené N (2013) A Structural Risk-Neutral Model for Pricing and Hedging Power Derivatives. Mathematical Finance 23:387–438

Alexander CO (1999) Correlation and cointegration in energy markets. In: RISK publications (ed) Managing Energy Price Risk, (2nd Edition), pp 291–304

Asche F, Osmundsen P, Sandsmark M. (2006) The UK Market for Natural Gas, Oil and Electricity: Are the Prices Decoupled. The Energy Journal 27:27–40

Bachmeiner LJ, Griffin JM (2006) Testing for market integration crude oil, coal, and natural gas. The Energy Journal 2:55–71

Bencivenga C, Sargenti G, D'Ecclesia RL (2010) Energy markets: crucial relationship between prices. In: Corazza M, Pizzi C (eds) Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer Verlag, Dordrecht Heidelberg London Milan New York, pp 23–32

Carmona R, Durrleman V (2003) Pricing and Hedging Spread Options. Society for Industrial and Applied Mathematics 45:627–685

Dempster MAH, Medova E, Tang K (2008) Long term spread option valuation and hedging. Journal of Banking & Finance 32:2530–2540

Diebold FX, Gunther T. A., Tay A. S. (1998) Evaluating Density Forecasts with Applications to Financial Risk Management. International Economic Review 34:868–883

Dixit AK, Pindyck RS (1994) Investment under uncertainty. Princeton University Press, Princeton, N.J

Eydeland A, Wolyniec K (2003) Energy and power risk management: New developments in modeling, pricing, and hedging. Wiley, Hoboken, N.J

Fama EF (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance 25:383–417

Felix BJ, Weber C (2012) Gas storage valuation applying numerically constructed recombining trees. European Journal of Operational Research 216:178–187

Frikha N, Lemaire V (2013) Joint Modelling of Gas and Electricity Spot Prices. Applied Mathematical Finance 20:69–93

Gardner D, Zhuang Y (2000) Valuation of Power Generation Assets: A Real Options Approach. Algo Research Quarterly 3:9–20

Gibson R, Schwartz E (1990) Stochastic convenience Yield and the Pricing of Oil Contingent Claims. The Journal of Finance 45:959–976

Gneiting T, Balabdaoui F, Raftery A (2007) Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69:243–268

Gondzio J, Kouwenberg R, Vorst T (2003) Hedging options under transaction costs and stochastic volatility. Journal of Economic Dynamics and Control 27:1045–1068

Graf C, Wozabal D (2013) Measuring competitiveness of the EPEX spot market for electricity. Energy Policy 2013:948–958

Guo X, Beskos A, Siddiqui A (2014) The natural hedge of a gas-fired power plant. Computational Management Science 13:63–86

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton, N.J.

Huisman R, Mahieu R (2003) Regime jumps in electricity prices. Energy Economics 25:425–434

Huisman R, Huurman C, Mahieu R (2007) Hourly electricity prices in day-ahead markets. Energy Economics 29:240–248

Hull JC (2011) Options, Futures, and Other Derivatives, 8th. Pearson College Div, Boston et al.

Janczura J, Weron R (2012) Goodness-of-fit testing for the marginal distribution of regime-switching models with an application to electricity spot prices. Advances in Statistical Analysis

Joets M, Mignon V (2011) On the link between forward energy prices: A nonlinear panel cointegration approach. Energy Economics:1170–1175

Kallabis T, Pape C, Weber C (2016) The plunge in German electricity futures prices: Analysis using a parsimonious fundamental model. Energy Policy 95:280–290

Kovacevic R, Paraschiv F (2014) Medium-term planning for thermal electricity production. OR Spectrum - Quantitative Approaches in Management 36:723–759

Kovacevic R, Wozabal D (2014) A semiparametric model for electricity spot prices. IIE Transactions Operations Engineering & Analytics 46

Kroese DP, Taimre T, Botev ZI (2011) Handbook of Monte Carlo methods. Wiley, Hoboken

Lloyd SP (1982) Least Squares Quantization in PCM. IEEE Transactions on Information Theory 28:129–137

Maciejowska K, Nowotarski J (2016) A hybrid model for GEFCom2014 probabilistic electricity price forecasting. International Journal of Forecasting 32:1051–1056

Mjelde JW, Bessler DA (2009) Market integration among electricity markets and their major fuel source markets. Energy Economics 31:482–491

Müsgens F (2006) Quantifying Market Power in the German Wholesale Electricity Market using a Dynamic Multi-Regional Dispatch Model. The Journal of Industrial Economics LIV:471–498

Panagiotidis T, Rutledge E (2007) Oil and gas markets in the UK: Evidence from a cointegrating approach. Energy Economics 29:329–347

Pape C, Hagemann S, Weber C (2016) Are Fundamentals Enough? Explaining Price Variations in the German Day-Ahead and Intraday Power Market. Energy Economics 54:376–387

Pape C, Vogler A, Woll O, Weber C (2017) Forecasting the distributions of hourly electricity prices: Accounting for serial correlation patterns and non-normality of price distribution. HEMF Working paper

Ripley BD (2009) Stochastic Simulation, 1st edn. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, NY

Rosenblatt M (1952) Remarks on a Multivariate Transformation. The Annals of Mathematical Statistic 23:470–472

Schwartz E (1997) The Stochastic Behavior of Commodity Prices: Implications for valuation and hedging. The Journal of Finance 52:929–973

Sunderkötter M, Weber C (2012) Valuing fuel diversification in power generation capacity planning. Energy Economics 34:1664–1674

Tseng C, Barz G (2002) Short-Term Generation Asset Valuation: A Real Options Approach. Operations Research 50:297–310

Weber C (2005) Uncertainty in the electric power industry: Methods and models for decision support. International series in operations research & management science, vol 77. Springer, New York

Weber C (2007) Strompreismodellierung - Berücksichtigung empirischer Verteilungen für nicht-speicherbare Güter am Beispiel Elektrizität. ESSENER UNIKATE - Berichte aus Forschung und Lehre 29:92–101

Weber C, Vogel P (2014) Contingent certificate allocation rules and incentives for power plant investment and disinvestment. Journal of Regulatory Economics 43:292–317

Weigt H, Hirschhausen C von (2008) Price formation and market power in the German wholesale electricity market in 2006. Energy Policy 36:4227–4234

Weron R (2014) Electricity price forecasting: A review of the state-of-the-art with a look into the future. International Journal of Forecasting 30:1030–1081

Woll O, Weber C (2011) Hybride Ansätze zur Preismodellierung im Kontext von Portfoliomanagement und Kraftwerksbewertung. In: VDI (ed) 9. Fachtagung Optimierung. VDI-Verl, Düsseldorf, pp 93–104

Wozabal D, Graf C, Hirschmann D (2014) The effect of intermittent renewables on the electricity price variance. OR Spectrum 38:687–709

Ziel F (2016) Forecasting Electricity Spot Prices Using Lasso: On Capturing the Autoregressive Intraday Structure. IEEE Transactions on Power Systems 31:4977–4987

Zugno M, Pinson P, Madsen H (2013) Impact of Wind Power Generation on European Cross-Border Power Flows. IEEE Transactions on Power Systems 28:3566–3575

Appendix A Additional material for modeling the short-term stochastics

Let $\boldsymbol{C}$ be the correlation matrix of $u_{d,h}$. The eigenvalues λ and corresponding eigenvectors $\mathbf{v}$ may be computed such that

$$\boldsymbol{Cv} = \lambda \boldsymbol{v} \quad (6\text{-}1)$$

A detailed description of PCA and the interpretation of the resulting eigenvalues λ and eigenvectors v can be found in .

In matrix notation, the relationship between the error term $u_{t,h}$(matrix $\mathbf{U}$) and factor $f_{t,i}$(matrix $\mathbf{F}$) is the following:

$$\mathbf{U} = \boldsymbol{VL}\mathbf{F} \quad (6\text{-}2)$$

Matrix $\mathbf{V}$ contains the eigenvectors $\boldsymbol{v}$ and the diagonal matrix $\mathbf{L}$ the square root of the eigenvalues λ.

$$\mathbf{U} = \left(u_{t,h}\right) \qquad \boldsymbol{F} = \left(f_{t,i}\right)$$

$$\boldsymbol{V} = (\boldsymbol{v_1 v_2} \ \dots \ \boldsymbol{v_{24}}) \qquad \boldsymbol{L} = \begin{pmatrix} \sqrt{\lambda_1} & 0\dots & 0 \\ 0 & \sqrt{\lambda_2}\dots & 0 \\ \vdots & \vdots\ddots & \vdots \\ 0 & 0\dots & \sqrt{\lambda_{24}} \end{pmatrix} \quad (6\text{-}3)$$

Eqs. (6-4) and (6-5) show the ARMA-GARCH specification for the factors $f_{t,i}$

$$f_{t,i} = \mu_i + \alpha_{1,i} f_{t-1,i} + \alpha_{2,i} \zeta_{t-1,i} + \zeta_{t,i} \ with \ \zeta_{t,i} \sim \mathrm{N}(0, \sigma_{t,i}) \quad (6\text{-}4)$$

$$\sigma_{t,i}^2 = \gamma_{0,i} + \gamma_{1,i}\, \sigma_{t-1.i}^2 + \gamma_{2,i}\, \zeta_{t-1,i}^2 \quad (6\text{-}5)$$

with the error $\zeta_{t,i}$ of the ARMA-part a and for the GARCH-part a constant term γ_0, the standard deviation from the previous day $\sigma_{t\text{-}1,i}^2$ and the quadratic error from the previous time step $\zeta_{t\text{-}1,i}^2$.

Appendix B Significance of the mean reversion parameters

Table 14 Significance of the mean reversion parameter for the Model 1

Newey-West (2009-2011)						
Delta Spread	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
spread	-.0064474	.0034977	-1.84	0.066	-.0133137	.0004189
_cons	.0419821	.0283656	1.48	0.139	-.013702	.0976663

The mean reversion parameter for the clean-spark-spread shows significance below the 10 percent level. For longer time spans and after 2011 the significance is below the 2.5 percent level but due to the financial crisis, we did not extend the estimation horizon beyond 2008/09.

Table 15 Significance of the mean reversion parameter for the Model 2

Newey-West (2009-2011)						
Delta Gas TTF	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
TTF	-.0040284	.0042373	-0.95	0.342	-.0123466	.0042898
_cons	.0883743	.0974907	0.91	0.365	-.1030083	.2797569
Newey-West (2009-2011)						
Delta Coal-API2	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Coal	-.0028628	.0053284	-0.54	0.591	-.013323	.0075974
_cons	.0348067	.0528995	0.66	0.511	-.0690397	.138653
Newey-West (2009-2011)						
Delta EUA CO2	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
CO2	-.0067506	.0069859	-0.97	0.334	-.0204645	.0069633
_cons	.080438	.1000731	0.80	0.422	-.1160141	.2768902

The GBM assumption is valid because the commodity prices show no significant mean reversion.

Appendix C Additional material for the valuation model

V_{t,T_1,T_2} equals the discounted expected profit from plant operation

$$V_{t,T_1,T_2} = \sum_{T=T_1}^{T_2} e^{r(T-t)} E\left[max(S_{n,T} - Cost_{n,T}, 0)\right], with\ t \leq T_1 \leq T_2 \quad (6\text{-}6)$$

$c_{n,T}$are the operational costs arising from producing electricity. $X_t^{total}$is the total price of inputs (fuel price $X_{f,t}$ including emissions $X_{CO_2,t}$ in EUR/MWh$_{th}$)

$$X_t^{total} = X_{f,t} + X_{CO_2,t} \cdot CO_{2_{f,u}}. \quad (6\text{-}7)$$

The cost function $C_{n,t}^{Op}$ depends on the fuel quantities $P_{n,t}^{Fu}$ and other costs c_{other}

$$Cost_{n,t}^{Op} = X_t^{total} \cdot P_{n,t}^{Fu} + Cost_{other} \cdot P_{n,t}^{El} \quad (6\text{-}8)$$

The fuel quantity is a linear function of the electricity output $P_{El,n}$.

$$P_{n,t}^{Fu} = a_0 O_{n,t} + a_1 \cdot P_{n,t}^{El}, \quad (6\text{-}9)$$

with a_1 the marginal heat rate when the power plant is operating, a_0 represents additional fuel consumption at the minimum stable operation level P_{min}. $O_{n,t}$ is a state variable indicating whether the plant is off or on $(0, 1)$. Oher equations for the constraints can be found in Tseng and Barz (2002) and Weber (2005). Start-up costs $C_{n,t}^{St}$ depend on the state variable $U_{s,t}$=$(0, 1)$, indicating a start-up.

$$Cost_t^{St} = \begin{cases} a_0^{St} + a_1^{St} \cdot X_t^{total}, if\ U_{n,t} = 1, \\ 0, \quad if\ U_{n,t} = 0 \end{cases}. \quad (6\text{-}10)$$

a_0^{St} includes labor plus fixed operating costs and maintenance expenses related to start-ups and a_1^{St} are the fuel costs for a start-up. The costs for staying online are the fuel costs during times with negative spreads. The optimal electricity output $P_{n,t}^{El}$ given the operation state $O_{n,t}$ is

$$P_{n,t}^{El}(O'_{n,t}) = O_{n,t}(O'_{n,t})\left(P_{min} + (P_{max} - P_{min}) \cdot \mathbb{1}_{S_{n,t} - a_1 X_t^{total}}\right) \quad (6\text{-}11)$$

To model minimum operation and minimum shutdown times, the state variable $O_{n,t}$ is generalized by the state variable $O'_{n,t}$ which counts the hours since start-up if positive or the hours since shutdown if negative. State transitions are only possible between certain values of $O'_{n,t}$. For example, a start-up is only possible once the shutdown time has reached the minimum shutdown time R_{down}, i.e., when $O'_{n,t}$=-R_{down}. A detailed description of these state constraints is given by Tseng and Barz (2002) and similarly by Gardner and Zhuang (2000).

Appendix D Point and distributional goodness-of-fit test

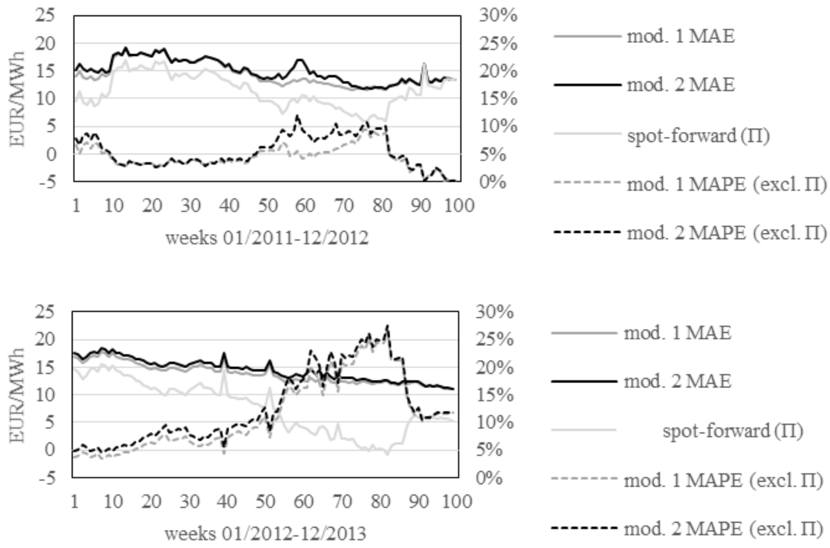

Fig. 7 Goodness-of-fit test: MAE and MAPE for Model 1 and Model 2 in 2012 and 2013

Appendix E Diebold-Mariano-test of equal predictive performance per hour

Table 16 Summary of Diebold-Mariano-test of equal predictive performance for separate hours 2012 and 2013

	2012	**2012**	**2012**	**2013**	**2013**	**2013**
Hour	Preferred Model 1	Preferred Model 2	Equal pred. (rej. H0)	Preferred Model 1	Preferred Model 2	Equal pred. (rej. H0)
1	80.0%	8.0%	12.0%	38.0%	30.0%	32.0%
2	74.0%	18.0%	8.0%	48.0%	24.0%	28.0%
3	72.0%	14.0%	14.0%	46.0%	20.0%	34.0%
4	68.0%	16.0%	16.0%	44.0%	18.0%	38.0%
5	76.0%	20.0%	4.0%	54.0%	28.0%	18.0%
6	78.0%	20.0%	2.0%	52.0%	36.0%	12.0%
7	74.0%	20.0%	6.0%	44.0%	40.0%	16.0%
8	62.0%	36.0%	2.0%	40.0%	48.0%	12.0%
9	56.0%	40.0%	4.0%	40.0%	48.0%	12.0%
10	54.0%	44.0%	2.0%	40.0%	48.0%	12.0%
11	30.0%	56.0%	14.0%	40.0%	48.0%	12.0%
12	22.0%	76.0%	2.0%	40.0%	48.0%	12.0%
13	22.0%	76.0%	2.0%	40.0%	48.0%	12.0%
14	22.0%	76.0%	2.0%	40.0%	48.0%	12.0%
15	22.0%	76.0%	2.0%	40.0%	48.0%	12.0%
16	22.0%	72.0%	6.0%	40.0%	46.0%	14.0%
17	42.0%	48.0%	10.0%	40.0%	48.0%	12.0%
18	52.0%	44.0%	4.0%	40.0%	48.0%	12.0%
19	58.0%	38.0%	4.0%	40.0%	48.0%	12.0%
20	60.0%	36.0%	4.0%	40.0%	48.0%	12.0%
21	58.0%	38.0%	4.0%	40.0%	48.0%	12.0%
22	30.0%	68.0%	2.0%	40.0%	48.0%	12.0%
23	26.0%	56.0%	18.0%	40.0%	48.0%	12.0%
24	44.0%	12.0%	44.0%	18.0%	6.0%	76.0%
Total	**50.0%**	**42.2%**	**7.8%**	**41.0%**	**40.3%**	**18.7%**

Appendix F Power plant value for delivery years 2012 and 2013

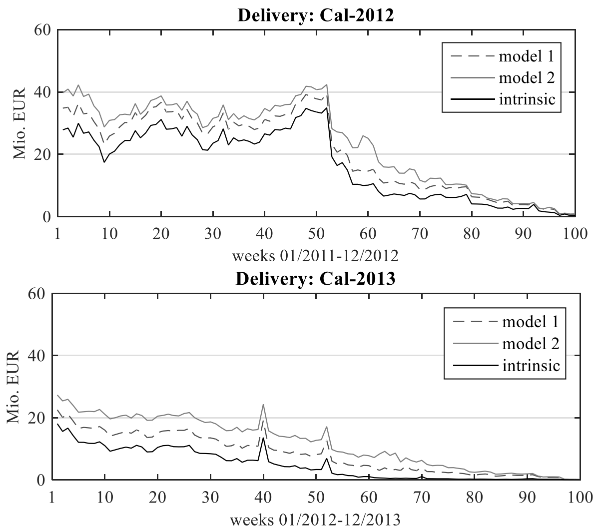

Fig. 8 Intrinsic ($\bar{V}$) and total (V) option value of the plant over time for delivery in 2012 (top) and 2013 (bottom)

The results for the monthly delivery periods are provided Fig. 9 and Fig. 10.

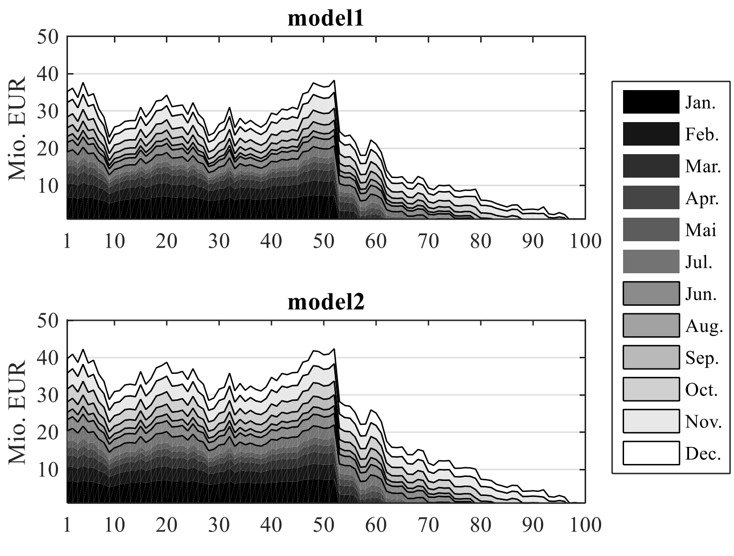

Fig. 9 Total option value (V) of single month over time for delivery in 2012

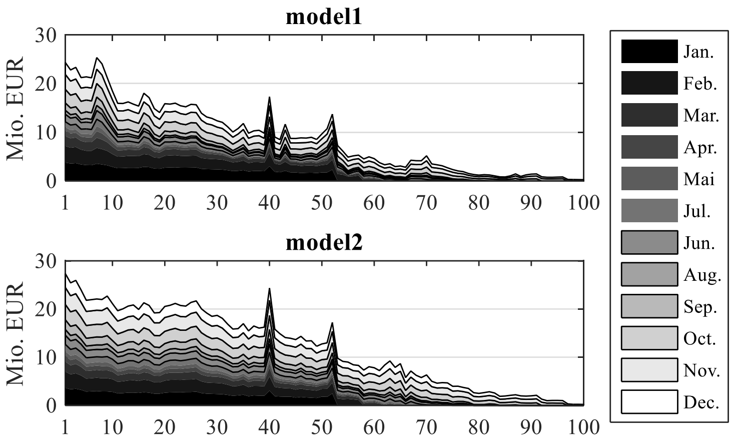

Fig. 10 Total option value (V) of single month over time for delivery in 2013

Appendix G Hedge amounts over time

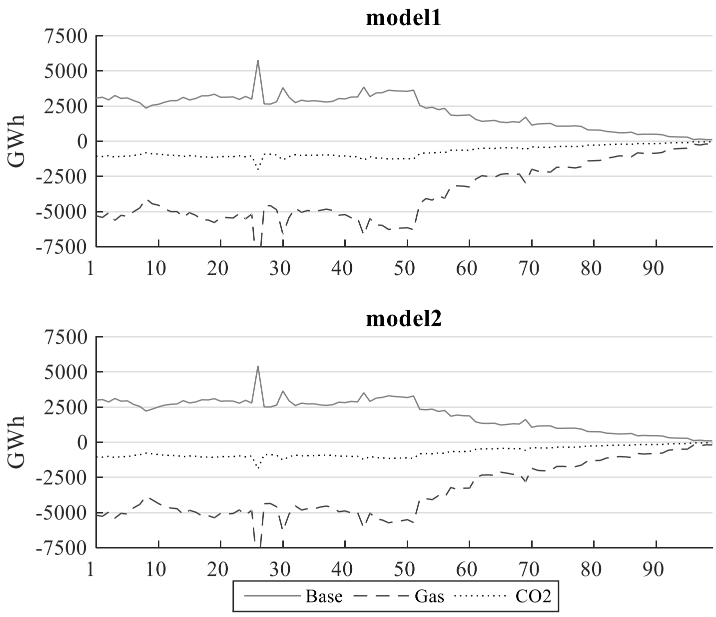

Fig. 11 Hedge amounts over time for delivery 2012

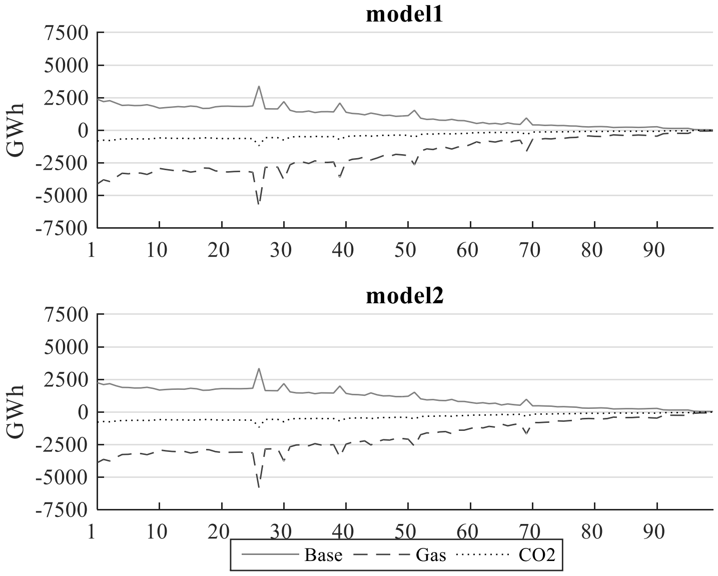

Fig. 12 Hedge amounts over time for delivery 2013

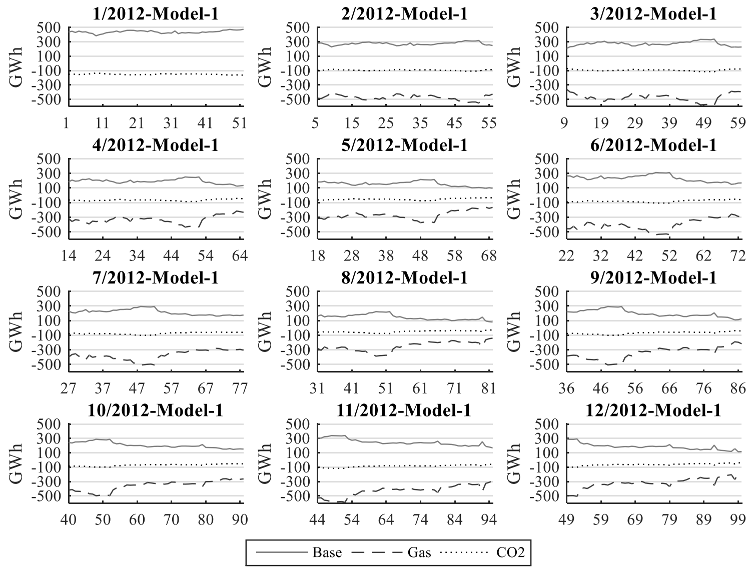

Fig. 13 Hedge amounts over time for Model 1 for monthly products of year 2012

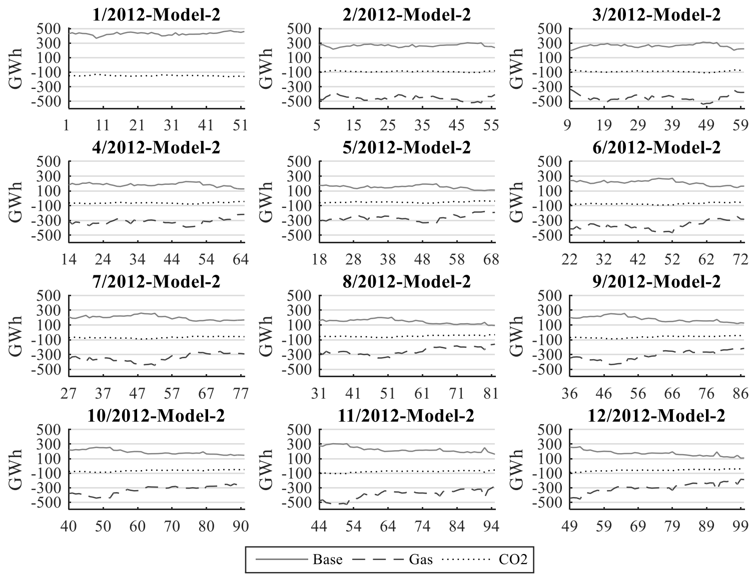

Fig. 14 Hedge amounts over time for Model 2 for monthly products of year 2012

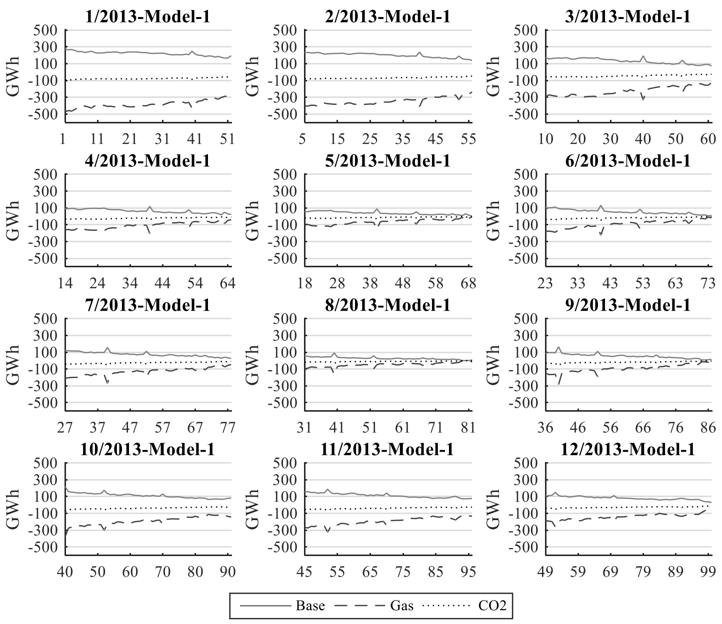

Fig. 15 Hedge amounts over time for Model 1 for monthly products of year 2013

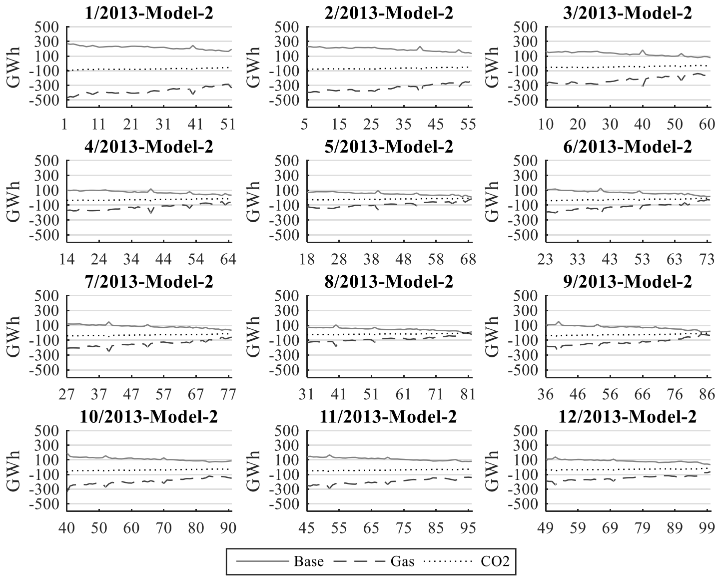

Fig. 16 Hedge amounts over time for Model 2 for monthly products of year 2013

Correspondence

Dr. Christian Pape
House of Energy Markets and Finance

Berliner Platz 6-8, 45127 Essen

Universitätsstr. 12, 45117

Tel. +49 201 183-2399
E-Mail christian.pape@uni-due.de

Dr. Oliver Woll

Centre of European Economic Research (ZEW)

L 7, 1, 68161 Mannheim, Germany

E-Mail woll@zew.de

Prof. Dr. Christoph Weber

Chairholder

House of Energy Markets and Finance

Berliner Platz 6-8, 45127 Essen

Universitätsstr. 12, 45117

Tel. +49 201 183-2399

Fax +49 201 183-2703

E-Mail web.hemf@wiwinf.uni-due.de

House of Energy Markets and Finance

University of Duisburg-Essen, Germany

Universitätsstr. 12, 45117, Essen

Tel. +49 201 183-2399

Fax +49 201 183-2703

E-Mail web.hemf@wiwinf.uni-due.de

Web www.hemf.net