The power of weather

Christian Huurman, Francesco Ravazzolo and Chen Zhou
The power of weather

Christian Huurman, Francesco Ravazzolo and Chen Zhou *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.
The power of weather*

Christian Huurman  Francesco Ravazzolo  Chen Zhou†

Financial Engineering  Norges Bank  De Nederlandsche
Associates  Bank

November 30, 2009

Abstract

This paper examines the predictive power of weather for electricity prices in day-
ahead markets in real time. We find that next-day weather forecasts improve the
forecast accuracy of Scandinavian day-ahead electricity prices substantially in
terms of point forecasts, suggesting that weather forecasts can price the weather
premium. This improvement strengthens the confidence in the forecasting model,
which results in high center-mass predictive densities. In density forecast, such
a predictive density may not accommodate forecasting uncertainty well. Our
density forecast analysis confirms this intuition by showing that incorporating
weather forecasts in density forecasting does not deliver better density forecast
performances.

Key words: Electricity prices, weather forecasts, point and density forecasts,
GARCH models.

JEL Classification Code: C53, G15, Q40.

---

*We are very gratefully to participants at 3rd International Conference on Computational and
Financial Econometrics (Cyprus 2009) for their helpful comments and we are in particular indebted
to Qaisar Farooq Akram, Jonas Andersson and Rob Hyndman for many useful comments. We thank
Nord Pool ASA for providing data. The views expressed in this paper are our own and do not
necessarily reflect the views of De Nederlandsche Bank and/or Norges Bank.

†Corresponding author. De Nederlandsche Bank, P.O. Box 98, 1000AB Amsterdam, The Nether-
lnds. Tel: +31 20 5243189. E-mail addresses: christian.huurman@fea.com (Christian Huurman),
francesco.ravazzolo@norges-bank.no (Francesco Ravazzolo), c.zhou@dnb.nl (Chen Zhou)
1 Introduction

This paper focuses on the predictive power of weather forecasts for electricity prices. The aim is to assess the performance in terms of both point and density forecasting of stochastic models with weather forecasts relative to the performance of models that ignore this information. In particular, we base our evaluation throughout different types of electricity prices (peak hours and daily average) and different markets.

Two decades ago, the electricity industry was strongly regulated and electricity prices reflected short-term production costs. Hence, back then the electricity price did not reflect temporal effects such as seasonality, weather and business activity. But this all changed when many governments worldwide started reforming their electricity industry as of the mid-1990s. Currently, the economic law of demand and supply determines the price in marketplaces where electricity can be traded in spot or forward (i.e. hour-ahead, day-ahead, month-ahead, or longer contracts). Trading electricity on liquid power wholesale markets, for example, the day-ahead markets\(^1\) which are by far the most liquid power wholesale markets, see Escribano, Pena, and Villaplana (2002), Lucia and Schwartz (2002) and Koopman, Ooms, and Carnero (2007), makes the price series predictable due to the aforementioned temporal effects.

Many studies have documented these stylized facts on in-sample analysis in electricity price models. Bunn and Karakatsani (2003) provide a thorough review of the stochastic price models presented in these studies and classify these into three groups: random walk models, basic mean-reversion models, and extended mean-reversion models that incorporate time-varying parameters (to control for seasonality and volatility patterns). They conclude that the idiosyncratic price structure has not been accurately described. Note that the results reported in these studies are often obtained from in-sample tests; hence they do not resolve the issue of the out-of-sample predictive value of power models.

While a few studies have recognized the need for modelling weather directly, these mainly addressed the weather effect on electricity sales (see, for example, the special issue of Journal of Econometrics 1979). Moral-Carcedo and Vicns-Otero (2005) study temperature effects on the variability of daily electricity demand in Spain, and document a non-linear relationship between variations in temperature and the demand response. Similar findings are described in Cancelo, Espasa, and Grafe (2008). More attention to the relationship between prices and weather is, however, given in very recent studies. Knittel and Roberts (2005) compare price models that incorporate seasonal and temperature variables with models that do not include these variables on hour-ahead electricity prices obtained from the California market, and provide preliminary evidence that the former models significantly outperform in terms of forecasting

\(^1\)On these markets, hourly prices are quoted for delivery of electricity at certain hours on the next day.
accuracy. Kosater (2006) conducts a similar exercise on the European Energy Exchange (the German market), but he extends the set of weather variables to wind velocity, to proxy for windmill producers, and uses non-linear models in the class of Markov regime-switching models. He finds that the significance of the weather-price relationship for forecasting is confined to certain hours. Huisman (2007) focuses on how temperature influences the probability of a spike and he shows that the difference between the actual and expected temperature significantly influences this probability.

As agents submit their bids and offers for delivery of electricity in all hours of the next day, weather conditions on the delivery day cannot be exactly known at the time when electricity prices are quoted. Therefore, we introduce weather forecasts, which are available in real time when prices are traded, in stochastic price models to forecast day-ahead prices in two bidding areas of the Nordic Power Exchange, Oslo and Eastern Denmark. We include a set of weather variables (temperature, precipitation and wind speed) to approximate the fact that both electricity demand and supply are subject to weather conditions and we implement specific models for different bidding areas due to the heterogeneity in weather conditions and production plants. To the best of our knowledge, a similar data set is not shared by any of the other papers focusing on electricity prices.

We find that an ARIMA model extended with power transformations of next-day weather forecasts yields the better point forecasting results in terms of root mean square prediction error for predicting day-ahead prices. In particular, this model outperforms, in terms of forecast accuracy, ARIMA models extended with actual weather on the delivered day, or extended with only determinist trends. We show that weather forecasts, even if subject to forecast errors, seem to have substantial explanatory power to anticipate ex-ante price jumps. Furthermore, we show that the relation between prices and weather forecasts is non-linear, and a model that incorporates weather forecasts must be carefully specified depending on data of interest.

As a further step we apply our models to density forecasting. Density forecasts are particular useful for spot market participants to hedge risk and for financial institutions to price derivatives on spot prices. As the utility functions of these players may differ, we stick to statistical measures. We follow Kitamura (2002), Mitchell and Hall (2005), Amisano and Giacomini (2007) and Kascha and Ravazzolo (2009), and apply the Kullback-Leibler divergence or Kullback-Leibler Information Criterion (KLIC) in our evaluation. Our results show that weather forecasts do not improve density forecast accuracy. Forecasting uncertainty, for example in terms of distributional shape, plays a key role in density forecasting. Standard assumptions such as normal approximation are therefore misleading and give losses even when the model fits data well, implying very difficult computational challenges for electricity derivative agents. More skewed predictive densities as in Panagiotelis and Smith (2008) could be considered.

The remainder of the paper is structured as follows. Section 2 introduces the day-ahead power markets and presents the data. Section 3 describes the forecasting models. Section 4 discusses the empirical results for point forecasts. Section 5 is devoted to density forecasting. Section 6 concludes.
2 Data

On January 1, 1991, the Norwegian government started a deregulation process for its electricity industry which resulted in the establishment of the first national power market for short-term delivery of power (real-time and day-ahead\(^2\)) in the world, the Nordic Power Exchange (NPX). Two years later, in 1993, the range of products was extended to include financial derivative contracts that have longer maturity horizons. A few years later, Sweden joined the NPX (1996), soon followed by Finland (1998), Western-Denmark (1999) and Eastern-Denmark (2000). Since 2003 all customers of Scandinavian electricity markets may trade freely in this market. The NPX, now also named Nord Pool ASA, is considered to be the most liquid wholesale market worldwide. Nord Pool ASA consists of a physical market, a financial market, and a trading emission allowances market. We only focus on the physical market, precisely the day-ahead spot market (Elspot). For more details on Nord Pool ASA we refer to NordPool (2004) and www.nordpool.com.

The Nord Pool market is largely dependent on electricity that is generated by renewable sources. In particular, hydro-plants, which use water stored in reservoirs or lakes, are dominant in Norway and partly Sweden; wind plants, which use wind to produce electricity, are dominant in Denmark. We again refer to NordPool (2004) for more details.

Electricity prices are affected by regional and temporal influences due to the transportation and transmission limits of electricity. This statement is particularly important in the Nord Pool market. For instance, when a power plant stops in the eastern part of Sweden this failure only affects the power supply in the surrounding region and does not affect the power supply in the western part of Sweden and the rest of the market. Similarly, rainfall in the southern part of Norway will potentially affect the regional demand and/or supply curve, but not the bidding curves in other regions. Nord Pool deals with the regional effects by splitting the market into several bidding and price areas. Therefore, we take into account the Nord Pool bidding area prices separately, rather than examining the Elspot system price (which is a weighted average of the bidding prices in all Nord Pool bidding areas). We examine two out of the eleven bidding areas in Nord Pool, i.e. the Oslo area and Eastern Denmark area. It is interesting to note that these areas are the most densely populated areas in Scandinavia.

2.1 Electricity prices

The dataset used in this study consists of day-ahead prices in EUR/MWh for Oslo and Eastern Denmark from the period December 24, 2003 to March 14, 2006\(^3\). Nord Pool

\(^2\)We recall from section 1 that day-ahead means that prices are quoted on day \(t\) for delivery of electricity at certain hours on the day \(t + 1\).

\(^3\)Electricity prices may be available for a longer sample, but weather forecasts are available to us only for this sample.
provides bidding area prices both in the local currency and in EUR. We choose EUR to compare directly the two area prices. Daily prices are computed as the arithmetic mean of the available 24 hourly price series on the physical market of each country. In our analysis we also use peak prices, which are defined as the average price of hourly prices from 8 am to 8 pm.

As in Wilkinson and Winsen (2002) and Lucia and Schwartz (2002) we start from a statistical analysis of the data we have. Figure 1 plots the time series, the log transformations and the first differences of log transformation of the daily day-ahead electricity prices; Table 1 reports some descriptive statistics. A first casual look discloses an erratic behavior of the prices. The series follow a small positive increasing trend with several spikes. Interestingly, prices in Oslo have more negative spikes than positive spikes. The wind-power drivers of the Danish market makes prices substantially different. The price level is higher; their distribution is non-normal; their volatility is very high as their kurtosis; and their skewness is positive. Oslo prices have a more regular distribution, but a Jarque-Bera test rejects the null hypothesis of normality for each of the six series. The series are characterized by weekly patterns. From Table 1 we can observe that prices are lower during the weekend than on working days. Moreover, as in Misiorek, Trueck, and Weron (2006) we find evidence of higher prices on Monday. Yearly patterns, well documented in other studies, are more difficult to identify probably due to our shorter sample. The logarithmic transformation reduces the spike behavior of the prices and makes moments of the distribution of electricity prices more similar to standard distributions, in particular for Eastern Denmark. We decide therefore to model log prices and present the analysis in function of them.

Electricity prices are very persistent and possibly close to non-stationary. Table 1 shows that the sample autocorrelations are high up to 14-day lags. The Dickey Fuller test on the series does not reject the null hypothesis of non-stationarity at any level of significance for Oslo prices and 1% level of significance for Eastern Denmark prices. This result prompts us to model the first difference prices, even if the Dickey Fuller test does not account for non-linear trends which may exist, but are not known to us. Modelling the first difference is also robust to potential structural breaks caused by the emergence of CO$_2$-Certificates in 2005. For example, electricity market participants have started to account for CO$_2$-costs when they calculate the marginal costs of coal-fired plants. Because quantifying the increase is difficult, the first difference seems to be a proper solution. From Figure 1, we can observe another stylized fact: volatility clustering. Dramatic spikes tend to occur in clusters, mainly as a result of consecutively exceeding the system capacity.

---

4We briefly discuss some stylized facts; we refer for a more detailed analysis, for example, to Lucia and Schwartz (2002) and Pilipovic (1997).

5We have also modelled the log level, but forecast results are less accurate, therefore we have decided to exclude it.
2.2 Weather forecasts

Our decision to focus on specific bidding areas of Nord Pool is motivated by the nature of weather variables. Weather observations and relative forecasts refer to a square area around the measurement station. Hence, it is impossible to have single observations that cover the entire countries of the markets under consideration. Nevertheless, the areas that we study are small and homogenous in terms of weather. Therefore, we use weather observations and forecasts for Oslo and Copenhagen.\(^6\) The weather around Oslo shall well approximate the weather in the area itself, the most populated zone of this bidding area, and to the south of Oslo along the seacoast where most of the electricity for south-east Norway is produced. The weather in the area of Copenhagen may be a proxy for the weather of the city of Copenhagen, again the most populated zone of this bidding area, and of Zealand, the main island in Eastern Denmark.

Daily average temperatures in degrees Celsius, total daily precipitation in mm, and daily average wind power in m/s are applied. Weather observations on day \(t\) and weather forecasts of the same day made at time \(t - 1\) are obtained from the EHAMFORE index, which is provided by Meteorlogix (www.meteorlogix.com).\(^7\) We assume that market operators use the weather forecasts provided by Meteorlogix in their decisions. This assumption is realistic considering the market share of Meteorlogix in providing real-time information services in the agricultural, energy, and commodity trading markets, and Bloomberg in providing data to operators.

Figures 2-3 plot actuals and forecasts of temperature, precipitation and wind in Oslo, and temperature and wind in Copenhagen.\(^8\) Temperatures are highly persistent, have highly seasonal patterns, and forecasts are quite precise. The correlation between temperature observations and their forecasts made the day before is higher than 0.9 for both areas. The wind speed is particularly high in Eastern Denmark, and forecasts are less accurate. The correlation between observations and forecasts is around 0.5 for both areas. We also notice that wind forecasts have a quite stable pattern in the initial months of 2004, because the Meteorological Institute applies a different scaling forecasting model in those months. We decide to keep these forecasts to extend the sample period as much as we can, because we believe that market operators had received this information in real time and used it to take their decisions. Actual precipitation and forecasted precipitation in the Oslo area are quite different. First, forecasting precipitation accurately is rather difficult; and second, observations are often zero, but models always forecast positive (possibly small) numbers.

Some graphical relations between the forecasted weather variables and electricity prices may be identified. For example, high precipitation in Oslo at the end of May

\(^6\)The combination of different stations could be applied, but data from minor cities are scarce.

\(^7\)Data from the EHAMFORE index are available in Bloomberg. Data on realized (or observed) precipitation on day \(t\) in Oslo from Meteorlogix are partially integrated with values of several different stations around Oslo from the Meteorological Institute (www.met.no) to fill in missing points

\(^8\)The graph of precipitation variable in Copenhagen is not reported as this variable is never chosen in the model selection procedure.
2004 or in October 2004 corresponds to low prices; a few days of very low temperatures in Oslo in February 2005 correspond to high prices; strong wind in Eastern Denmark at the end of 2004 and beginning of 2006 is associated with low prices. Using real weather does not change the conclusion: a graphical analysis is not satisfactory because the relation between weather variables and electricity prices is possibly non-linear as we will discuss in section 4.1.

Several studies, see e.g. Koopman, Ooms, and Carnero (2007) and Deng (2004), argue that the water reservoir is the key variable to plan production, even more than the amount of precipitation. Even if we agree with this view, we emphasize that in this paper we work with local prices, and local water reservoirs is in most cases not a public information. This is particularly true when the number of electricity producers is high, as for example in the Oslo area. And we do not yet have a model to construct regional water reservoirs. Therefore, we opt to apply total precipitation as a proxy for water reservoirs.

3 Forecasting models

Knittel and Roberts (2005) show that traditional time series approaches such as ARMA and ARIMA models provide more accurate results in forecasting electricity prices than their continuous counterparts. Starting from these findings we construct several models that may cope with the stylized facts of electricity prices.

3.1 ARIMA

ARIMA$_1$ The first model is a traditional time series approach to model electricity prices, the autoregressive moving average (ARMA) model (see, for example, Hamilton (1994)). The ARMA(p, q) model implies that the current value of the investigated process (say, the log price) $P_t$ is expressed linearly in terms of its past p values (autoregressive part) and in terms of the q previous values of the process $\epsilon_t$ (moving average part). The ARMA modelling approaches assume that the time series under study is (weakly) stationary. If it is not, a transformation of the series to stationarity is necessary, such as first order differentiating. The resulting model is known as the autoregressive integrated moving-average model (ARIMA). We define the ARIMA$_1$ model as:

$$\phi(L)p_t = \theta(L)\epsilon_t,$$

where $p_t = (P_t - P_{t-1})$, $\phi(L)$ and $\theta(L)$ are the autoregressive and moving average polynomials in the lag operator $L$ respectively, defined as:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p,$$
$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - ... - \theta_q L^q,$$

and $\epsilon_t$ is an independent and identically distributed (iid) noise process with zero mean and finite variance $\sigma^2$. 

7
**ARIMA**  ARIMA models apply information related to the past of the process and do not use information contained in other pertinent time series. However, as the data analysis shows, electricity prices are generally governed by various fundamental factors, such as seasonality and load profiles. Following the analysis in the previous section, we extend the ARIMA as:

\[
\phi(L)(p_t - X_t) = \theta(L)\epsilon_t,
\]

where \(X_t = \sum_{i=1}^{k} \psi_i x_{i,t} \) is the \((k \times 1)\) vector of dummies at time \(t\) and \(\psi = (\psi_1, \psi_2, ..., \psi_k)'\) is a \((k \times 1)\) vector of coefficients. We consider two variables: a dummy with values 0 on working days and 1 on holidays, and a dummy with values 1 on Monday and 0 elsewhere. In the empirical application, we define it as ARIMA\(_2\) model and take it as our benchmark.

**ARIMAX**  Adverse weather conditions may change the demand for electricity, and may also affect production. Low levels of precipitation and/or wind speed may reduce the supply of energy, in particular in electricity markets which depend on renewable producer plants, such as Norway and Denmark. Furthermore, producer plants may study future weather conditions to estimate demand and plan their supply optimally.

Forecasts of the average daily temperature in degrees Celsius, precipitation in mm and wind speed in m/s are applied as further explanatory variables in the ARIMA\(_2\) model. The ARIMAX model is:

\[
\phi(L)(p_t - X_t - W_t) = \theta(L)\epsilon_t,
\]

where \(W_t = \sum_{j=1}^{l} \phi_j w_{j,t} \) is the \((l \times 1)\) vector of weather forecast variables at time \(t\), and \(\phi = (\phi_1, \phi_2, ..., \phi_l)'\) is a \((l \times 1)\) vector of coefficients. This model includes deterministic components that account for genuine regularities in the behavior of electricity prices and stochastic components that come from weather shocks.

Knittel and Roberts (2005) apply a similar model for forecasting California electricity prices, where the set of weather variables consists of the level, as well as the square and the cubic of realized temperature. As in Knittel and Roberts (2005), we allow non-linearity in the relation between prices and weather variables by including the level, the square and the cubic of the temperature forecasts. We also add a variable to measure wind speed since wind may play a role both in the demand for electricity in such markets where electricity is used for heating - people associate stronger wind with colder temperature - and in the supply of wind power plants. And, finally, we introduce precipitation as further explanatory variable to approximate the supply in hydro dominated plants. Again, we allow non-linearity by using the level and the square of the precipitation and wind forecasts. Precipitation and wind forecasts are always positive and therefore we do not consider it useful to include their cubic transformations.

**ARIMAX-A**  Our weather forecasts for day \(t\) differ from actual weather on day \(t\), in particular for total precipitation, as we discuss in section 2. We investigate whether these differences have an impact on electricity price forecasts by replacing in the previous model weather forecasts made at time \(t - 1\) for weather on day \(t\) with
actuals on day $t$. In other words, we assume that when making the price decision, it is possible to produce a completely accurate forecast on the weather of the next day. This is in fact not achievable in practice; but we would like to investigate whether increasing the forecast accuracy on weather may lead to a better forecast on electricity prices. We define this model ARIMAX-A.

### 3.2 GARCH models

**ARIMA-GARCH** ARIMA models assume homoscedasticity, i.e. constant variance and covariance functions, but the preliminary data analysis has disclosed that electricity prices exhibit volatility clustering. We extend previous models by assuming a time-varying conditional variance of the noise term. The heteroskedasticity is modelled by a generalized autoregressive conditional heteroskedastic GARCH($r, s$) model (Bollerslev (1986)). Relaxing the assumption of homoscedasticity may change the parameter estimates of ARIMA, and consequently the out-of-sample forecast of the investigated process. Furthermore, GARCH models produce volatility forecasts which may improve accuracy of density forecasts.

The model is:

$$
\phi(L)(p_t - X_t) = \theta(L)\epsilon_t \tag{6}
$$

$$
\epsilon_t = \nu_t h_t^{1/2} \quad \text{with} \quad h_t = \alpha_0 + \sum_{i=1}^{s} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{r} \beta_j h_{t-j}, \tag{7}
$$

where $\nu_t$ is an independent and identically distributed (iid) noise process with zero mean and variance 1, and the coefficients have to satisfy $\alpha_i \geq 0$ for $1 \leq i \leq s$, $\beta_j \geq 0$ for $1 \leq j \leq r$, and $\alpha_0 > 0$ to ensure that the conditional variance is strictly positive.$^9$

**ARIMAX-GARCH** The ARIMAX model can also be extended by assuming a noise process with a time-varying conditional variance, that is:

$$
\phi(L)(p_t - X_t - W_t) = \theta(L)\epsilon_t \tag{8}
$$

$$
\epsilon_t = \nu_t h_t^{1/2} \quad \text{with} \quad h_t = \alpha_0 + \sum_{i=1}^{s} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{r} \beta_j h_{t-j}. \tag{9}
$$

**ARIMAX-GARCHX** Koopman, Ooms, and Carnero (2007) find that seasonal factors and other fixed effects in the variance equation are also important to estimate electricity prices. We follow these suggestions and assume the conditional variance of the noise term in ARIMAX model to be time-varying and modelled with an extended GARCH expression. The model is specified as:

$$
\phi(L)(p_t - X_t - W_t) = \theta(L)\epsilon_t \tag{10}
$$

$^9$We only report results for the ARIMA$_2$ model with GARCH residuals, but not for the ARIMA$_1$ model.
\[ \epsilon_t = \nu_t h_t^{1/2} \quad \text{with} \quad h_t = \alpha_0 + \sum_{i=1}^s \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^r \beta_j h_{t-j} + \sum_{f=1}^{k+l} \varrho_f z_{f,t}, \]

where \( z_t = [x'_t, w'_t]' \), and \( \varrho = (\varrho_1, \varrho_2, ..., \varrho_{k+l})' \) is a \((k + l) \times 1\) vector of coefficients. Since Koopman, Ooms, and Carnero (2007) assume autoregressive fractionally integrated moving average noises, we do consider integration of order one to simplify the estimation. Koopman, Ooms, and Carnero (2007) include water reservoirs, which are excluded in this study for the aforementioned reasons.

4 Empirical Results

We apply the models described in section 3 to our dataset, and firstly assess which model performs best in terms of point forecasting accuracy. We start by estimating the set of models using the complete sample to have an ex-post predictability idea.

We apply a non-linear ordinary least square (NLS) estimator (Davidson and MacKinnon (1993)) for ARIMA-type models and the quasi maximum likelihood (QML) estimator (Davidson and MacKinnon (1993) and Greene (1993)) for GARCH family models.

We restrict our ARIMA-type models to be ARIMA(7,0). Autocorrelation analysis and in-sample criteria would suggest more complex ARIMA forms. In particular, lags 14, 21, 28 and 56 still have high correlations, suggesting longer weekly and monthly price behaviors. However, the risk of over-parametrization and the evidence presented in previous studies, for example Lucia and Schwartz (2002), show that an ARIMA(7,0) specification provides optimal forecasts on daily day-ahead electricity prices. MA parameters are often not significant. Following the same reasoning we choose a GARCH(1,1) specification for the time-varying variance models.

The inclusion of the weather variables follows from statistical evidence. We allow different transformations of the weather forecast variables on the two markets as the weather may affect the supply of electricity, which is different in the two markets. We minimize the Akaike information criterion to specify the model.

4.1 In-sample analysis

The in-sample analysis is based on the overall sample, from December 24, 2003 to March 14, 2006. Table 2 reports estimation results. We focus in the discussion on Oslo prices. The coefficient estimate of the variable \( D_{\text{hol},t} \), which is a dummy variable with value 1 if day \( t \) is not a working day or 0 if day \( t \) is a working day, is statistically significant and negative. The coefficient estimate of the variable \( D_{\text{Monday},t} \), which is a dummy variable with value 1 if day \( t \) is Monday and 0 elsewhere, is statistically significant and positive. These results confirm evidence in Table 1 and tell us some important information on how agents quote prices. Prices are set lower during weekends (or not a working day) when electricity utilization is lower, but when the market is back to standard working conditions, agents seem initially to overreact and then adjust their positions.
Even if the fit is quite high, residuals of the ARIMA$_2$ model are not normally distributed. We think that weather variables can explain part of them. Figure 4 shows that the errors of the ARIMA$_2$ model have a non-linear relation with the daily average temperature and the total precipitation, but have a linear-like relation with the wind speed. In the ARIMAX for temperature forecasts we use the level, the square and the cubic; for precipitation and wind forecasts we use the level and the square. The selection criterion results in the following specific set of weather variables:

$$W_t = a_1 \text{Temp}_t + a_2 \text{Temp}_t^2 + a_3 \text{Temp}_t^3 + b_1 \text{Prec}_t + b_2 \text{Prec}_t^2 + \gamma \text{Wind}_t,$$

where $\text{Temp}_t$, $\text{Prec}_t$ and $\text{Wind}_t$ are the forecasts of daily average temperature, total precipitation and wind speed, respectively, on day $t$. The square of the wind is then excluded in the selection procedure. Figure 4 indicates a direct linear relation between prices and wind speed and the empirical findings are consistent with the graphical analysis.

The interpretation of estimated coefficients for the weather variables is not often straightforward due to non-linear relations. For example, the temperature forecasts affect the day-ahead electricity price via the following function:

$$f(\text{Temp}_t) = a_1 \text{Temp}_t + a_2 \text{Temp}_t^2 + a_3 \text{Temp}_t^3.$$  

Taking the first order derivative, we get

$$\frac{df(\text{Temp}_t)}{d\text{Temp}_t} = a_1 + 2a_2 \text{Temp}_t + 3a_3 \text{Temp}_t^2.$$  

By substituting in the previous equation $a_1$, $a_2$ and $a_3$ with their empirical estimates, and solving $\frac{df(\text{Temp}_t)}{d\text{Temp}_t} = 0$, we find the roots are 4, 19.5. When temperature is lower than 4 or higher than 19.5, the derivative is negative; when temperature is in the interval [4, 19.5], the derivative is positive. We first focus on the switching point 4. From Figure 2, we observe that 4 degree is lying in the middle of the range of temperature forecasts. When the forecast is below 4 degree, it has a negative impact on the electricity price. When above 4 degrees, its impact turns to be positive. Intuitively, this reflects that when the temperature forecast is relative higher or lower than the switching point 4 degrees, the consumption of electricity will rise. Meanwhile the difficulty in producing electricity increases when it is extremely cold or hot. The other switch point 19.5 lies around the upper bound of the temperature forecasts. In fact, only in 1.6% cases in our sample, the temperature forecast is higher than 19.5 degree. Considering the estimation uncertainly, it is likely that such a switch point does not have economic impact. Furthermore, a close look at the estimated coefficients $a_1$, $a_2$ and $a_3$ shows that $a_3$ is not significantly different from zero$^{10}$. By omitting the last item in the

$^{10}$We decide to keep the cubic transformation of temperature among predictors in the equation even if its coefficient is not statically different from zero because it minimizes the Akaike criterion. It is possible that this variable may influence electricity prices only in some marginal and specific parts of their domain, but still worthwhile to consider.
derivative of \( f \) function, we obtain a single switch point as \(-a_1/(2a_2) = 3.8\). The switching point around 4 degree does exist even in this case, while the other switching point does not appear in the relation between temperature forecasts and day-ahead electricity price.

Following similar reasoning, we find that precipitation influences prices negatively such that higher precipitation leads to lower prices. This is consistent with the hydropower nature of Oslo electricity markets. The coefficient of the wind speed is not statically different from zero, but excluding it does not minimize the selection criterion. The estimate is negative and we do not have strong economic explanations for it. We note that the Akaike criterion is maximized when weather variables are introduced, supporting this strategy.

### 4.2 Out-of-sample analysis

We forecast day-ahead Oslo log electricity prices from January 1, 2005 to March 14, 2006. We repeat the selection procedure in section 4.1 over the initial in-sample period, from December 24, 2003 to December 31, 2004. The reduced specific form in this shorter sample of all models remains the same as in full sample analysis. This indicates that relations do not seem to be sample dependent. In forecasting, when new values are available we re-estimate models to produce a new forecast, but we do not re-specify it. An expanding window is used, which means that to forecast the price of day \( t \), all the previous data are included.

We compare point forecasts from different models using the Root Mean Square Prediction Error (RMSPE), defined as

\[
RMSP E = \sqrt{\frac{1}{n} \sum_{s=1}^{n} (P_{T+s} - \hat{P}_{T+s})^2},
\]

where \( P_{T+s} \) is the log price at time \( T + s \), \( \hat{P}_{T+s} \) is the forecasted log price at time \( T + s \), and \( n = 438 \) is the number of days being forecasted. Results are given in Table 3. Weather forecasts improve point forecasts. The ARIMAX model provides the most accurate forecasts. The improvement with respect to the ARIMA\(_2\) model is around 6% in term of RMSPE. We compare whether the difference is significant by applying the Diebold-Mariano test (Diebold and Mariano (1995)). The null hypothesis is \( H_0 \): The square of the forecast errors are equal, and we reject it. \(^{11}\)

\(^{11}\)The limit approximation of the Diebold-Mariano test is not valid for nested models, which is also our case. As, for example, Busetti, Marcucci, and Veronese (2009) discuss, the test rejects with a lower probability the null hypothesis in the case of nested models. McCracken (2007) and Clark and West (2007), for example, propose alternative methods. For the sake of brevity, we report only the Diebold-Mariano test as our findings in favor of the ARIMAX should not change since the null of equal predictability is already rejected.
Figure 5 can help to interpret this evidence. Figure 5 shows the 60-day average RMSPE for the ARIMA\textsubscript{2} model and the ARIMAX model. From the graph, we find that, when the error of the former model is at a relatively lower level, the errors of the two models are similar, but when there is a higher error due to possible jumps or outliers, the latter model often predicts better. Price jumps are mainly due to problems of inelasticity of demand, and of non-storability of electricity with consequent shortages in the supply. These problems often arise when weather conditions are adverse.

Adding GARCH specification to approximate evidence of volatility clustering does not provide more accurate results. In particular, the ARIMAX-GARCHX proposed by Koopman, Ooms, and Carnero (2007) does not outperform the benchmark model. Our ARIMAX-GARCHX model is, however, simpler than the estimated model of Koopman, Ooms, and Carnero (2007) and the set of variables is different. Finally, augmenting ARIMA models with actual weather does not help. The ARIMAX-A model gives similar results to the ARIMA\textsubscript{2} model, but the ARIMAX outperforms it. This seems to confirm that market agents do not have information on tomorrow’s actual weather when they take decisions, so the information they can eventually use to settle electricity prices is the forecasts of tomorrow’s weather. In other words, the forecasting power of the weather variables on the electricity prices is contained in the weather forecasts rather than in the actual weather. Increasing forecast accuracy on weather does not improve electricity forecasting if the information is not somewhat public available.

We test the robustness to electricity peak hour prices. Peak hour prices are defined as the average of the hourly prices from 8 am to 8 pm. As a matter of fact prices are more likely to be volatile in that specific frame. The models for peak hour prices specified from general to specific using the Akaike selection criterion are the same as for daily prices. We could anticipate this, considering the role of peak hours in determining daily prices. Forecasting results are reported in Table 3. The ARIMAX is clearly the best model, which seems to confirm our conjecture that weather forecasts help to anticipate sharp movements in prices.

We implement our analysis in a different Nord Pool bidding area, Eastern Denmark. We re-estimate and re-specify all the models. We find that the variable precipitation is not significant, which can be explained by the fact that Denmark does not rely on hydropower. We develop a forecasting exercise similar to the Oslo study. Again the ARIMAX model provides the best forecasts. The improvement is higher than 10%, and it is statistically significant at the 1% level. Figure 5 shows that also for the Eastern Denmark sample weather forecasts help to partially anticipate price jumps. Furthermore, from Table 3 we can observe that using actual weather reduces forecast accuracy compared to using weather forecasts and results for the ARIMAX model are again robust to peak hour prices.

5 Density forecasting

Following deregulation, electricity derivatives have also been introduced in electricity markets worldwide to allow energy utilities to protect themselves against different risks. For example, financial derivatives are used extensively by market participants to hedge
their risk exposure to price fluctuations. Electricity derivatives are used by financial institutions as another source of investment opportunities. The evaluation of electricity derivatives is particularly difficult due to the specific features of the underlying asset. For example, electricity cannot be short sold and its trading is not continuous, making the Black-Scholes derivative valuation imprecise. Alternative valuation methods are often applied to estimate the density of the payoff from a electricity derivative (and to derive the fair price of the contract or Value-at-Risk figures). All these valuation methods are based at least on estimates of the first two (or more) moments of the underlying asset. Therefore we decide to focus on density forecasting as we think that models that provide more accurate density forecasts should be preferred in derivatives pricing. First we present how we compute and evaluate density forecasts. Then we discuss empirical evidence.

5.1 Computing & Evaluating Density Forecasts

We compute 1-step ahead density forecasts using a normal approximation. Assuming that the past errors and coefficients are known, the conditional expectation corresponds to the point forecast of each individual model. The forecast variance is computed differently for ARIMA models and GARCH models. In the former class of models we approximate the forecast error variance with the in-sample estimate of the error variance $\sigma^2$ in section 3.1. In the latter class we use the conditional time-varying variance $h_t$ in section 3.2. The predictive density given by any of the models in the suite is then

$$f_{t+1,i}(y_{t+1}) \sim N(\mu_{t+1,i}, \sigma_{t+1,i}^2),$$

(12)

where $\mu_{t+1,i}$ is the point forecast for model $i$, $\sigma_{t+1,i}^2$ is the forecast variance for model $i$ both made at time $t$ for time $t + 1$, and $i = 1, \ldots, 7$ as in section 3.

The essential problem in evaluating density forecasts is that the true density is never observed - not even after the random variable is drawn. Additional difficulties arise, if one wants to compare multiple models that are misspecified and possible nested, see Timmermann (2006) for a discussion.

We follow Kitamura (2002), Mitchell and Hall (2005), Amisano and Giacomini (2007) and Kascha and Ravazzolo (2009), and use the Kullback-Leibler divergence or Kullback-Leibler Information Criterion (KLIC) to evaluate our results. The KLIC chooses the model which on average gives higher probability to events that have actually occurred. Specifically, the KLIC distance between the true density $f_t$ of a random variable $y_t$ and some candidate density $f_{t,i}$ obtained from model $i$ is defined as

$$\text{KLIC}_t = \int f_t(y_t) \ln \frac{f_t(y_t)}{f_{t,i}(y_t)} dy_t,$$

(13)

Under some regularity conditions, a consistent estimate can obtained from the average of the sample information, $y_1, \ldots, y_T$, on $f_t$ and $f_{t,i}$:

$$\overline{\text{KLIC}} = \frac{1}{T} \sum_{t=1}^{T} \ln \frac{f_t(y_t)}{f_{t,i}(y_t)}.$$

(14)
Even though we do not know the true density, we can still compare multiple densities, \( f_{t,i} \). For the comparison of two competing models, it is sufficient to consider only the latter term in the above sum,

\[
-\frac{1}{T} \sum_{t=1}^{T} \ln f_{t,i}(y_t),
\]

for all \( i \) and to choose the model for which the expression in (15) is minimal.

Differences in KLIC can be statistically tested. We apply the test of equal accuracy of two density forecasts developed by Mitchell and Hall (2005). Suppose there are two density forecasts, \( f_{t,1}(y_t) \) and \( f_{t,2}(y_t) \), and consider the loss differential

\[
d_t = \ln(f_{t,1}(y_t)) - \ln(f_{t,2}(y_t)).
\]

The null hypothesis of equal accuracy is then:

\[
H_0 : E(d_t) = 0.
\]

The sample mean \( \bar{d}_t = \frac{1}{T} \sum_{t=1}^{T} d_t \) has, under appropriate assumptions, the limiting distribution:

\[
\sqrt{T} (\bar{d}_t - d_t) \rightarrow N(0, \Omega).
\]

This is the departure point for testing the null hypothesis.

The KLIC and its based test are relative measures, which compare the forecasting accuracy between two models. Different from that, we also consider a measure and a test on the absolute accuracy of the forecasting model. We compute \( PITS \), probability integral transforms, as in, for example, Little, McSharry, and Taylor (2008) and apply the Berkowitz (2001) test for zero mean, unit variance and independence of the \( PITS \). Under the null hypothesis that the forecasting model is well calibrated, its \( PITS \) are uniformly distributed. Then, the Berkowitz (2001) likelihood ratio test follows a chi-square distribution with degrees of freedom 3, \( \chi^2(3) \). Therefore, tests based on the \( PITS \) provide an indication of whether the null of no calibration failure can be rejected.

### 5.2 Results

The results of the out-of-sample density forecasting are summarized in Tables 5-4. The table reports the out-of-sample forecasting performance of the individual models for daily log prices and peak hour log prices of the two countries. We compare the logarithmic scores to investigate relative performance and to find which is the most accurate model in terms of density forecasting. We repeat that higher logarithmic scores mean better performance. The results in Table 4 are different than in the previous section: the weather improves point forecast accuracy (in terms of RMSPE) but does not help in density forecasting (in terms of \( \ln S \)). The simple ARIMA\(_1\) model gives the best statistics for both applications with Oslo data and very close to the best one for Eastern Denmark daily prices. The differences from the ARIMA\(_2\) model are
statistically significant. The models with weather variables perform worse than the ARIMA$_1$ and ARIMA$_2$ models. How we construct density forecasts can explain these findings *ex-post*. We approximate the forecast variance for model $i$ by the in-sample residual variance. We describe in Table 2 and in section 4 that the ARIMAX model fits data better and forecasts more accurately in terms of point values than the other ARIMA specifications. This implies a smaller variance for the residuals, or in other words a smaller uncertainty for the density forecast of the ARIMAX. Figure 6 shows the patterns of the three variance forecasts. Trends are similar but the value for the ARIMAX model is substantially lower as the model is more precise in relative terms to capture shifts or jumps. The model is, however, imprecise in absolute terms and cannot capture the complete magnitude of the jumps. Even if it can partially anticipate these jumps compared to other specifications, a higher center-mass density forecast added to the normality assumption provides smaller scores in the tails. The lnS measure heavily punishes these realizations, favoring less precise density forecasts but with some higher probabilities in small, often rare tail events. The forecast uncertainty could be dealt with using, for example, a Bayesian approach as in Panagiotelis and Smith (2008). This is outside the scope of this paper.

The findings for the GARCH models are partially different. GARCH models result in very negative lnS statistics for Oslo, but the best ones for Eastern Denmark. The GARCH volatility, therefore, seems to approximate well Eastern Denmark prices and thus improve density forecasts. Adding weather variables to the GARCH equations does not change results.

The PITS test shows the goodness of calibration. Table 5 finds that for Oslo, all the models are rejected under 10% confidence level except the ARIMAX model. Incorporating weather forecast information in density forecasting generates the best model calibration, but as we discussed above it does not increase the predictive power for density forecasting. The result is robust when forecasting the density of peak hour electricity prices. For Eastern Denmark, only ARIMAX-GARCHX model is not rejected. Again the weather forecasts improve calibration, but only if associated to a more sophisticated structure on volatility forecasting. This agrees with the observation on the volatility clustering property for Eastern Denmark area.

### 6 Conclusion

In this paper we study whether weather variables have predictive power for electricity prices in real time. We collect forecasts of several weather variables for two bidding areas of Nord Pool, Oslo and Eastern Denmark, and we use them to predict the respective prices and thus improve density forecasts. Our empirical results suggest that weather forecasts play a central role in forecasting day-ahead prices. Weather forecasts give relevant information to partially anticipate spikes in prices, providing more accurate point forecasts compared to several alternative approaches. Results are robust to using actual weather and to peak hour price applications. We find that the relation between electricity prices and weather forecasts is highly non-linear and depends on the price drivers behind each bidding area. When we move to density forecasting, results are the
opposite. The fact that weather forecasts produces more accurate point forecasts for electricity prices strengthens our confidence on the model which results in higher center-mass predictive densities. When extreme (possibly unpredictable) events happen, such assumption has serious implications, deteriorating substantially the density forecasting performance.

There are several topics for further research. First, the set of weather forecasts might include other weather-related variables, such as water reservoir levels. Observations of these variables are not available to us, but results support the construction of specific models to compute approximations and forecasts of them. Second, considering the strong non-linear relation of prices and weather forecasts, weather forecasts might be included in other non-linear models, such as Markov regime-switching models, jump models or time-varying models, see Karakatsani and Bunn (2008). These specifications may cope better with jumps in forecasting exercises, and keep a more realistic value of uncertainty in their predictive densities compared to what we do in this study.
References


Deng, D., 2004, Modelling and investigating the relationship between electricity prices and water reservoir content at Nord Pool power exchange, Discussion paper, Gothenburg University.


Kosater, P., 2006, On the impact of weather on German hourly power prices, Discussion paper, University of Cologne.


Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Oslo</th>
<th>Eastern Denmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level</td>
<td>Log</td>
</tr>
<tr>
<td>Mean</td>
<td>30.43</td>
<td>3.403</td>
</tr>
<tr>
<td>St dev</td>
<td>4.965</td>
<td>0.154</td>
</tr>
<tr>
<td>Max</td>
<td>52.45</td>
<td>3.960</td>
</tr>
<tr>
<td>Min</td>
<td>17.16</td>
<td>2.843</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.319</td>
<td>0.513</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>6.278</td>
<td>4.768</td>
</tr>
<tr>
<td>Working days</td>
<td>30.90</td>
<td>3.420</td>
</tr>
<tr>
<td>No working days</td>
<td>29.35</td>
<td>3.366</td>
</tr>
<tr>
<td>Mondays</td>
<td>31.04</td>
<td>3.424</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.938</td>
<td>0.928</td>
</tr>
<tr>
<td>$\rho_7$</td>
<td>0.824</td>
<td>0.78</td>
</tr>
<tr>
<td>$\rho_{14}$</td>
<td>0.678</td>
<td>0.677</td>
</tr>
<tr>
<td>DF</td>
<td>-0.535</td>
<td>-1.678</td>
</tr>
</tbody>
</table>

Note: The table reports descriptive statistics on level and logarithm of electricity prices, and of logarithm first difference in Oslo and Eastern Denmark. Lines Mondays, working days and no working days give the sample average prices on Mondays, working days and no working days (weekends and holidays) respectively. Lines $\rho_1$, $\rho_7$ and $\rho_{14}$ give the 1st, 7th and 14th sample autocorrelation. The last line refers to the Augmented Dickey-Fuller test; one, two or three asterisks denote significance relative to the asymptotic null hypothesis respectively at 10%, 5% and 1% levels.
<table>
<thead>
<tr>
<th></th>
<th>Oslo</th>
<th>Eastern Denmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARIMA₁</td>
<td>ARIMA₂</td>
</tr>
<tr>
<td>( c )</td>
<td>0.001</td>
<td>-0.004</td>
</tr>
<tr>
<td>( \phi₁ )</td>
<td>-0.254</td>
<td>-0.215</td>
</tr>
<tr>
<td></td>
<td>[0.034]</td>
<td>[0.035]</td>
</tr>
<tr>
<td>( \phi₂ )</td>
<td>-0.237</td>
<td>-0.189</td>
</tr>
<tr>
<td></td>
<td>[0.035]</td>
<td>[0.036]</td>
</tr>
<tr>
<td>( \phi₃ )</td>
<td>-0.208</td>
<td>-0.162</td>
</tr>
<tr>
<td></td>
<td>[0.036]</td>
<td>[0.036]</td>
</tr>
<tr>
<td>( \phi₄ )</td>
<td>-0.240</td>
<td>-0.189</td>
</tr>
<tr>
<td></td>
<td>[0.036]</td>
<td>[0.036]</td>
</tr>
<tr>
<td>( \phi₅ )</td>
<td>-0.191</td>
<td>-0.124</td>
</tr>
<tr>
<td></td>
<td>[0.036]</td>
<td>[0.036]</td>
</tr>
<tr>
<td>( \phi₆ )</td>
<td>-0.076</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.034]</td>
<td>[0.038]</td>
</tr>
<tr>
<td>( \phi₇ )</td>
<td>0.240</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>[0.035]</td>
<td>[0.035]</td>
</tr>
<tr>
<td>( d₀ )</td>
<td>-</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.004]</td>
</tr>
<tr>
<td>( dₐ )</td>
<td>-</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>( a₁ )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.0002]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>( a₂ )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[2.48E-05]</td>
<td>[2.16E-04]</td>
</tr>
<tr>
<td>( a₃ )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[1.52E-06]</td>
<td>[8.42E-06]</td>
</tr>
<tr>
<td>( b₁ )</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[0.007]</td>
<td>[0.007]</td>
</tr>
</tbody>
</table>

**Note:** The table reports the coefficient estimates (and their standard errors between square brackets), and selection criteria tests of the models on Oslo log prices for models in section 3.
<table>
<thead>
<tr>
<th>Model</th>
<th>Oslo 24h</th>
<th>Oslo Peak</th>
<th>Eastern Denmark 24h</th>
<th>Eastern Denmark Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA_2</td>
<td>0.048</td>
<td>0.059</td>
<td>0.170</td>
<td>0.176</td>
</tr>
<tr>
<td>ARIMA_1</td>
<td>1.118</td>
<td>1.126</td>
<td>1.101</td>
<td>1.123</td>
</tr>
<tr>
<td>ARIMAX</td>
<td>0.942**</td>
<td>0.907**</td>
<td>0.905**</td>
<td>0.899**</td>
</tr>
<tr>
<td>ARIMAX-A</td>
<td>0.988</td>
<td>0.999</td>
<td>0.998</td>
<td>1.002</td>
</tr>
<tr>
<td>ARIMA-GARCH</td>
<td>1.029</td>
<td>1.009</td>
<td>1.028</td>
<td>0.984</td>
</tr>
<tr>
<td>ARIMAX-GARCH</td>
<td>1.010</td>
<td>0.996</td>
<td>1.035</td>
<td>1.019</td>
</tr>
<tr>
<td>ARIMAX-GARCHX</td>
<td>1.008</td>
<td>1.002</td>
<td>1.027</td>
<td>1.013</td>
</tr>
</tbody>
</table>

Note: The table reports forecasting statistics of the alternative models in Oslo and Eastern Denmark electricity markets using daily average prices (24h) or peak hours (8am-8pm), (Peak) prices. The first line in the table reports the value of RMSPE for the ARIMA_2 model while all the other lines report statistics relative to the ARIMA_2 model. Bold numbers indicate the best performing model in terms of RMSPE whereas asterisks indicate results of the forecast accuracy comparison t-type statistics as in Diebold and Mariano (1995) of the given models against those of the ARIMA_2 model. The null hypothesis is that the two forecasts have the same mean square error. One asterisk denotes significance relative to the asymptotic null hypothesis at 5%, two asterisks denote significance relative to the asymptotic null hypothesis at 1%.
Table 4: Out-of-sample forecasting results: Logarithmic Score

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Oslo 24h</th>
<th>Oslo Peak</th>
<th>Eastern Denmark 24h</th>
<th>Eastern Denmark Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA_1</td>
<td>-16.76**</td>
<td>-15.90**</td>
<td>-6.14**</td>
<td>-3.76**</td>
</tr>
<tr>
<td>ARIMA_2</td>
<td>-20.87</td>
<td>-20.52</td>
<td>-7.65</td>
<td>-4.67</td>
</tr>
<tr>
<td>ARIMAX</td>
<td>-22.42**</td>
<td>-21.98**</td>
<td>-8.01**</td>
<td>-4.95**</td>
</tr>
<tr>
<td>ARIMA-GARCH</td>
<td>-46.13**</td>
<td>-43.66**</td>
<td><strong>-6.08</strong></td>
<td>-0.68**</td>
</tr>
<tr>
<td>ARIMA-GARCHX</td>
<td>-53.15**</td>
<td>-43.92**</td>
<td>-6.48</td>
<td>-0.70**</td>
</tr>
<tr>
<td>ARIMA-GARCHX</td>
<td>-57.35**</td>
<td>-47.33**</td>
<td>-6.41</td>
<td>-0.73**</td>
</tr>
</tbody>
</table>

Note: The table reports logarithmic Score statistics for Oslo and Eastern Denmark electricity density forecasts using daily average prices (24h) or peak hours (8am-8pm), (Peak) prices. Bold numbers indicate the best performing model in term of logarithmic score, whereas asterisks indicate results of the forecast density comparison test as in Mitchell and Hall (2005) of the given models against those of the ARIMA_2 model. One asterisk denotes significance relative to the asymptotic null hypothesis at 5%, two asterisks denote significance relative to the asymptotic null hypothesis at 1%.
Table 5: Out-of-sample forecasting results: PITS

<table>
<thead>
<tr>
<th>Model</th>
<th>Oslo 24h</th>
<th>Oslo Peak</th>
<th>Eastern Denmark 24h</th>
<th>Eastern Denmark Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.059</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMAX</td>
<td>0.168</td>
<td>0.109</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMAX-A</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMA-GARCH</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMAX-GARCH</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>ARIMAX-GARCHX</td>
<td>0.088</td>
<td>0.000</td>
<td>0.197</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: The table reports the Likelihood Ratio p-value of the Berkowitz (2001) test of zero mean, unit variance and independence of the inverse normal cumulative distribution function transformed \( PITS \), with a maintained assumption of normality for transformed \( PITS \) for Oslo and Eastern Denmark electricity density forecasts using daily average prices (24h) or peak hours (8am-8pm), (Peak) prices. Numbers lower than 0.05 and 0.01 correspond to reject the null hypothesis that the forecasting model is well calibrated respectively at 5% and 1% level of significance respectively.
Note: The graphs in this figure present in Panel a) prices, in Panel b) log prices, and in Panel c) first differences log prices for Oslo (in the left panel) and for Eastern Denmark (in the right panel).
Figure 2: Weather variables: Oslo

Note: The graphs present in Panel a) actual (ACT) and forecasted (FOR) daily average temperature (in the left panel), and total precipitation (in the right panel); in Panel b) actual (ACT) and forecasts (FOR) on wind speed in the Oslo area.

Figure 3: Weather variables: Eastern Denmark

Note: The graphs in Figure 3 present actual (ACT) and forecasted (FOR) daily average temperature (in the left panel) and actual (ACT) and forecasted (FOR) wind speed in the Copenhagen area.
Figure 4: Scatter plot: Oslo

(a)

(b)

Note: The graphs in this figure present in Panel a) the scatter plot of the errors of ARIMA$_2$ model against the forecasts of the daily average temperature (in the left panel), and total precipitation (in the right panel); in Panel b) the scatter plot of the errors of ARIMA$_2$ model against the forecasts of the wind speed in Oslo.
Figure 5: 60-day average RMSPE

Note: The graphs in this figure present in Panel a) the 60-day moving average RMSPE given the ARIMA$_2$ and ARIMAX models in forecasting Oslo log electricity prices; in Panel b) the 60-day moving average RMSPE given the ARIMA$_2$ and ARIMAX models in forecasting Eastern Denmark log electricity prices.

Figure 6: Forecast standard deviation

Note: The graphs in this figure present the forecast standard deviation for the ARIMA$_1$, ARIMA$_2$ and ARIMAX models in density forecasting daily Oslo log electricity prices (in the left panel) and Eastern Denmark log electricity prices (in the right panel).
Previous DNB Working Papers in 2009

No. 198 Peter ter Berg, Unification of the Fréchet and Weibull Distribution
No. 199 Ronald Heijmans, Simulations in the Dutch interbank payment system: A sensitivity analysis
No. 200 Itai Agur, What Institutional Structure for the Lender of Last Resort?
No. 201 Iman van Lelyveld, Franka Liedorp and Manuel Kampman, An empirical assessment of reinsurance risk
No. 202 Kerstin Bernoth and Andreas Pick, Forecasting the fragility of the banking and insurance sector
No. 203 Maria Demertzis, The ‘Wisdom of the Crowds’ and Public Policy
No. 204 Wouter den Haan and Vincent Sterk, The comovement between household loans and real activity
No. 205 Gus Garita and Chen Zhou, Can Open Capital Markets Help Avoid Currency Crises?
No. 206 Frederick van der Ploeg and Steven Poelhekke, The Volatility Curse: Revisiting the Paradox of Plenty
No. 207 M. Hashem Pesaran and Adreas Pick, Forecasting Random Walks under Drift Instability
No. 208 Zsolt Darvas, Monetary Transmission in three Central European Economies: Evidence from Time-Varying Coefficient Vector Autoregressions
No. 209 Steven Poelhekke, Human Capital and Employment Growth in German Metropolitan Areas: New Evidence
No. 210 Vincent Sterk, Credit Frictions and the Comovement between Durable and Non-durable Consumption
No. 211 Jan de Dreu and Jacob Bikker, Pension fund sophistication and investment policy
No. 212 Jakob de Haan and David-Jan Jansen, The communication policy of the European Central Bank: An overview of the first decade
No. 213 Itai Agur, Regulatory Competition and Bank Risk Taking
No. 214 John Lewis, Fiscal policy in Central and Eastern Europe with real time data: Cyclicality, inertia and the role of EU accession
No. 215 Jacob Bikker, An extended gravity model with substitution applied to international trade
No. 216 Arie Kapteyn and Federica Teppa, Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice
No. 217 Mark Mink and Jochen Mierauf, Measuring Stock Market Contagion with an Application to the Sub-prime Crisis
No. 218 Michael Biggs, Thomas Mayer and Andreas Pick, Credit and economic recovery
No. 219 Chen Zhou, Dependence structure of risk factors and diversification effects
No. 220 W.L. Heeringa and A.L. Bovenberg, Stabilizing pay-as-you-go pension schemes in the face of rising longevity and falling fertility: an application to the Netherlands
No. 221 Nicole Jonker and Anneke Kosse, The impact of survey design on research outcomes: A case study of seven pilots measuring cash usage in the Netherlands
No. 222 Gabriele Galati, Steven Poelhekke and Chen Zhou, Did the crisis affect inflation expectations?
No. 223 Jacob Bikker, Dirk Broeders, David Hollanders and Eduard Ponds, Pension funds’ asset allocation and participants age: a test of the life-cycle model
No. 224 Stijn Claessens and Neeltje van Horen, Being a Foreigner among Domestic Banks: Asset or Liability?
No. 225 Jacob Bikker, Sherrill Shaffer and Laura Spierdijk, Assessing Competition with the Panzar-Rosse Model: The Role of Scale, Costs, and Equilibrium
No. 226 Jan Marc Berk and Beata K. Bierut, Communication in a monetary policy committee: a note
No. 227 Dirk Broeders, An Chen and Birgit Koos, An institutional evaluation of pension funds and life insurance companies
No. 228 Michael Hurd, Maarten van Rooij and Joachim Winter, Stock market expectations of Dutch households
No. 229 Maria Demertzis, Massimiliano Marcellino and Nicola Viegri, Anchors for Inflation Expectations
No. 230 Jan Willem van den End and Mostafa Tabbae, When liquidity risk becomes a macro-prudential issue: Empirical evidence of bank behaviour
No. 231 Maarten van Rooij, Annamaria Lusardi and Rob Alessi, Financial Literacy and Retirement Planning in the Netherlands
No. 232 Chen Zhou, Are banks too big to fail?
No. 233 Frederick van der Ploeg, Steven Poelhekke, THE PUNGENT SMELL OF “RED HERRINGS”: Subsoil assets, rents, volatility and the resource curse
No. 234 Mark Mink, Is Contagion in the Eye of the Beholder?
No. 235 Carin van der Cruijsen and Maria Demertzis, How Anchored Are Inflation Expectations in EMU Countries?
Financial acceleration of booms and busts