Lunina, Veronika

Working Paper

Joint Modelling of Power Price, Temperature, and Hydrological Balance with a View towards Scenario Analysis

Working Paper, No. 2016:30

Provided in Cooperation with:
Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Lunina, Veronika (2016) : Joint Modelling of Power Price, Temperature, and Hydrological Balance with a View towards Scenario Analysis, Working Paper, No. 2016:30, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:
https://hdl.handle.net/10419/260202

Working Paper 2016:30

Department of Economics
School of Economics and Management

Joint Modelling of Power Price, Temperature, and Hydrological Balance with a View towards Scenario Analysis

Veronika Lunina

November 2016

LUND UNIVERSITY

Joint Modelling of Power Price, Temperature, and Hydrological Balance with a View towards Scenario Analysis

Veronika Lunina

Knut Wicksell Centre for Financial Studies and Department of Economics, School of Economics and Management, Lund University, S-22007 Lund, Sweden

Abstract

This study presents a model for the joint dynamics of power price, temperature, and hydrological balance, with a view towards scenario analysis. Temperature is a major demand-side factor affecting power prices, while hydrobalance is a major supply-side factor in power markets dominated by hydrological generation, such as the Nordic market. Our time series modelling approach coupled with the skew-Student distribution allows for interrelations in both mean and volatility, and accommodates most of the discovered empirical features, such as periodic patterns and long memory. We find that in the Nordic market, the relationship between temperature and power price is driven by the demand for heating, while the cooling effect during summer months does not exist. Hydrobalance, on the other hand, negatively affects power prices throughout the year. We demonstrate how the proposed model can be used to generate a variety of joint temperature/hydrobalance scenarios and analyse the implications for power price.

Keywords:
spot power price, temperature, hydrological scenarios, VARFIMA-BEKK, skew-Student
JEL: C32, G17, Q41

1. Introduction

In this study, we develop a model for the joint evolution of the spot electrical power price, outdoor temperature, and hydrological balance. The model is relevant

Email address: nika.lunina@gmail.com (Veronika Lunina)

for power markets with a large share of hydrological generation, such as the Nordic market, and offers a wide range of opportunities for scenario analysis.

Consider the following example illustrating the fundamental relationship between power price, temperature, and hydrobalance. Energy producers and retailers plan their business activities based on the estimated demand for power (load) for a certain time horizon. During the heating seasons, the load is driven to a large extent by temperature. If a particular season is actually colder than expected, more power than planned will be consumed for heating purposes. Temperature affects demand, but to understand the implications for prices, we also need to consider the supply side of price formation. We define hydrobalance as the measure of the potential capacity of a hydrological power generation system. If a year has been relatively wet, with lots of precipitation, i.e., hydrobalance is high, this excess demand may be covered at a low cost without moving the price. On the contrary, a combination of low temperature and low hydrobalance is a major source of price risk in power markets. Therefore, it is natural to model these three variables as a system.

Both temperature modelling and hydrological modelling are large research areas on their own. Among the literature related to power markets, Halldin (2005) studies modelling of the time series of water inflows and stochastic optimization of a hydropower system. Green (2015) shows that the intra-daily profiles of the Nord Pool system price are affected by hydrological balance and develops an hourly forward curve model with hydrological dependence. Bivariate power-temperature models have been developed in Benth et al. (2012) and Caporin et al. (2012), for the purpose of pricing an exotic type of weather derivatives called energy quanto options. This study adopts an econometric modelling approach, similar in certain respects to that of Caporin et al. (2012), introducing the third dimension of hydrobalance into the system.

We analyse both the univariate properties of our three data series, and the dependencies between them in detail. The model is identified within the Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA) framework, coupled with a time-varying covariance process of Baba-Engle-Kraft-Kroner (BEKK) type. The need for fractional integration is motivated by long memory, observed in all series.

In addition to the stochastic part, our model contains a deterministic component capturing the yearly periodic patterns in power prices and temperatures. Due to the highly pronounced non-normal statistical properties of our data, we apply a flexible multivariate skew-Student distribution proposed in Bauwens and Laurent (2002, 2005), while treating the normal distribution as the benchmark for comparison. The model allows for interrelations, both in means and volatilities, restricted such that power price can be affected by temperature and hydrobalance, but not the other way around.

We find that in the Nordic market, the relationship between temperature and power price is driven by the demand for heating, while the cooling effect during summer months does not exist, likely due to mild temperature conditions. Hydrobalance, on the other hand, has a significant inverse effect on power prices throughout the year. Further, estimation results indicate the existence of volatility spillover effects from hydrobalance and temperature to power. Correlations between power and temperature show seasonal patterns, ranging from -0.5 during winter periods to 0 during summer periods. Correlations between power and hydrobalance oscillate around an average level of -0.25. Finally, the simulation exercise reveals the benefits of skew-Student distribution in reproducing the distinct non-Gaussian properties of both power price and meteorological series.

The remainder of the study is organized as follows. Section 2 dscribes the data and the results of preliminary data analysis. Section 3 presents the modelling framework, as well as the identification and estimation procedure. The empirical results are discussed in Section 4. Section 5 addresses the simulation from the model and provides a scenario analysis example, which demonstrates how our model can be utilized to generate a variety of joint temperature/hydrobalance scenarios and analysing the implications for power prices. Section 6 contains a summary and concluding remarks.

2. Data and preliminary analysis

This section describes the data and the results of preliminary data analysis, which lays the foundation for our choice of modelling framework. We first investigate the

properties of the univariate series that we want the model to reproduce, and then discuss the desired dependence structure.

2.1. The dataset

Our dataset consists of daily observations of the power price series, the temperature series, and the hydrobalance series. Since the model is designed specifically for power markets with a large share of hydrological generation, we use Nordic market data.[1] The sample period spans from January 1, 2008, to February 21, 2016.

We obtain the power price data from the Nordic power exchange, the Nord Pool. The Nord Pool Elspot is the spot market where agents trade power on an auction basis for physical delivery during each hour of the following day, which is why it is often referred to as the day-ahead market. The daily spot power price is the arithmetic average across 24 hourly prices, and is quoted in EUR/MWh. Being the common marketplace for the Nordic (Denmark, Finland, Norway, and Sweden) and the Baltic (Estonia, Latvia, and Lithuania) countries, Nord Pool is divided into a number of bidding areas, which can have different prices in the presence of transmission constraints. In addition to these area prices, all participating countries share a common *system spot price*, calculated under the assumption of unconstrained transmission capacity. In this study, we consider the Nord Pool system spot power price since it is the reference price for trading and clearing the majority of financial contracts.

The temperature and hydrobalance data are obtained from Thomson Reuters. We use the daily average temperature (DAT) in Sweden, which is the population-weighted average across a basket of several cities. Alternatively, we could consider the average temperature across all Nord Pool area countries, but since the Baltic countries joined the market during 2010 – 2013, it is more straightforward to use a single country as a proxy for the whole region. This does not lead to any loss of generality because temperature series in the individual Nordic and Baltic countries

[1] Hydro power is the largest generation source in terms of installed capacity in the Nordic power market. According to the Nord Pool, in a year with normal precipitation, hydro power accounts for half of Nordic countries' demand (98% in Norway, in particular).

are highly correlated, and Sweden would have had the largest population weight in the index anyway. The daily average temperature is the average of the minimum and the maximum temperature during a given day measured in degrees of Celsius.

Finally, the hydrobalance series represents the deviations of the total hydrological resources from the seasonal normal level measured in terms of energy capacity (TWh). The total hydrological resources are defined as the sum of the water reservoir content, the snow pack, and soil water, and reflect the available capacity of the hydro-power system. The seasonal normal levels of the hydrological resources for each day of a year are computed by Thomson Reuters based on the 1981 – 2005 period, and thus account for recent weather trends. Consequently, a positive (negative) hydrobalance value on a given day indicates that the hydrological condition is wetter (drier) than it has been on average for the same day of a year during 1981 – 2005. Most hydrological data are usually available at a weekly granularity, but Thomson Reuters provides daily Nordic hydrobalance series starting from 2008. This series contains more fundamental information than if we were to interpolate between the weekly observations, which motivates our choice of the sample period start.

2.2. Data analysis

Figure 1 plots the power price series along with the fitted seasonal mean function given by Eq. (1) below.

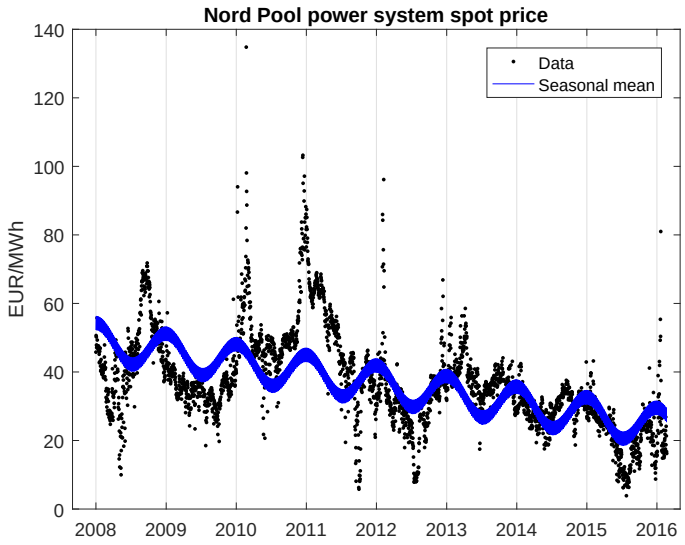


Figure 1: Time series of Nord Pool power system spot prices.

We can clearly see a decreasing trend, as an increasing amount of low marginal cost wind, solar, and biomass generation displaces coal and gas technologies. Another important feature is yearly seasonal patterns with higher prices during winter periods related to demand for heating and shorter day-light length. Seasonal patterns can be captured by a periodic function, such as sine or cosine. In addition, day-of-the-week effects in power prices have been widely documented, with lower prices observed on weekends and holidays due to limited business activity (see, e.g., Lucia and Schwartz, 2002). Therefore, we choose the following specification for power seasonal mean at time t, denoted $\Lambda_{1,t}$:

$$\Lambda_{1,t} = \lambda_{1,1} + \lambda_{1,2}\cos(\frac{2\pi}{365}(t - \lambda_{1,3})) + \lambda_{1,4}t + \lambda_{1,5}D_t, \tag{1}$$

where $\lambda_{1,i}$, $i = 1, ..., 5$, are the parameters to estimate, t is time measured in days, and D_t is a dummy variable taking a value of 1 if day t is a non-business day (i.e., a weekend or a holiday), and 0 otherwise. $\lambda_{1,1}$ represents the overall (non-seasonal) average price level, $\lambda_{1,2}$ is the amplitude of the mean price, and $\lambda_{1,3}$ is the phase angle. The amplitude of a cosine wave reflects how large the distance between peaks and troughs is. A phase angle shifts the time to adjust for the fact that yearly

maximum and minimum mean prices do not necessarily have to occur on January 1 and July 1, respectively. Note that the period of oscillation is equal to one year, or 365 days, ignoring leap years.

A closer look at Figure 1 reveals that power prices can have large upward spikes, followed by fast mean reversion. Spikes typically occur during winter seasons if high demand for power coincides with an unexpected supply-side shock, like an outage at a major power plant. It is worth noting that spikes are relatively less dramatic in power markets with a large share of hydro generation, such as the Nordic market, since water reservoirs can serve as a safety buffer against unforeseen imbalances between supply and demand. This, of course, heavily depends on hydrological conditions and whether there is enough excess capacity in the hydro-power system to provide this sort of safety cushion.

The temperature series is plotted in Figure 2, along with the fitted seasonal mean function given by Eq. (2) below.

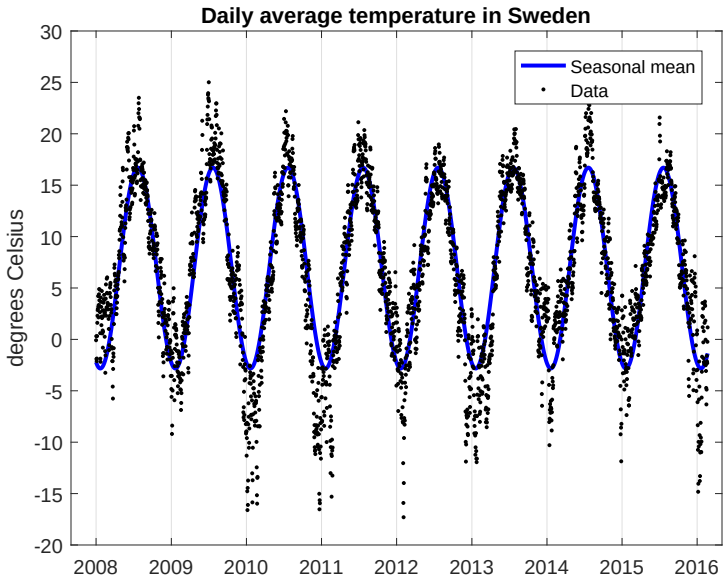


Figure 2: Time series of daily average temperature in Sweden.

We adopt the following specification for temperature seasonal mean at time t, denoted $\Lambda_{2,t}$:

$$\Lambda_{2,t} = \lambda_{2,1} + \lambda_{2,2}\cos(\frac{2\pi}{365}(t - \lambda_{2,3})), \tag{2}$$

where $\lambda_{2,1}$ is the overall (non-seasonal) average level of the series, $\lambda_{2,2}$ is the amplitude of the mean temperature, and $\lambda_{2,3}$ is the phase angle. Alaton et al. (2002) document a small yet statistically significant increasing linear trend in Stockholm temperatures during 1957 – 1997, which is attributed to global warming and urbanization. We do not find such an effect in our sample, probably since we are looking at a much shorter time span. Since our model is not meant for multi-year forecasts, we omitted the linear trend term. Benth and Šaltytė-Benth (2005) reach the same conclusion while examining the Norwegian temperature data during 1990 – 2003.

We form de-seasonalised power price and de-seasonalised temperature series by subtracting from the original series Λ_1 and Λ_2, respectively. Further analysis in this section is concerned with the properties of de-seasonalised series.

Finally, we examine hydrobalance. Note that hydrobalance is a series of deviations from the normal state, and therefore should have a long-run mean level of zero, assuming that the chosen normal state is representative of recent dynamics, or stable over long periods of time. However, when looking at shorter horizons, the sample mean can move away from zero, as is the case in our sample. As Figure 3 shows, dry periods are more common than wet periods during the sample.

We can also see how extremely persistent hydrological condition is: once a trend is established, it might take months for hydrobalance to revert back to zero. It is also fairly uncommon for the series to change signs within a single year; therefore, it makes sense to classify the whole years as 'wet' or 'dry'. Nevertheless, it is by no means binary, and a wide range of possible scenarios are likely. The hydrobalance series in itself is not seasonal, although the total hydrological resources do exhibit strong yearly seasonal patterns, with water reservoirs being gradually filled after the spring flood and melting of the snow pack.

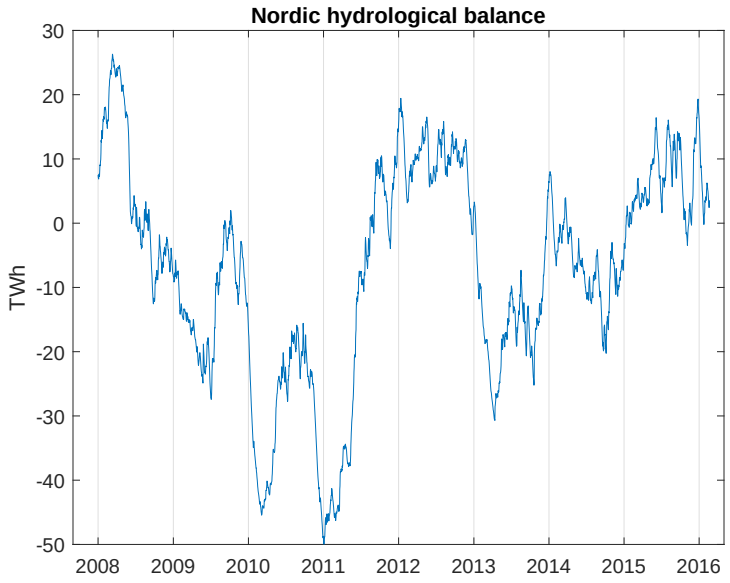


Figure 3: Time series of daily Nordic hydrological balance.

Table 1 contains the results of stationarity tests.

Table 1: Stationarity tests.

	Power price	Power log-price	Temperature	Hydrobalance
ADF	-5.33^{***}	-6.48^{***}	-14.35^{***}	-2.27
KPSS	0.39^{*}	0.33	0.36^{*}	0.81^{***}
Sample size	2974	2974	2974	2974

Note: The table reports the results of stationarity tests. ADF refers to the Augmented Dickey-Fuller test with the null hypothesis of a unit root. KPSS refers to Kwiatkowski et al.'s (1992) test with the null of a stationary I(0) process. Stationarity tests for power price, log-price, and temperature were applied to de-seasonalised series. The number of lags in the ADF tests were selected based on Schwartz Information Criterion. KPSS tests used the Bartlett kernel with Newey-West automatic bandwidth selection. Superscripts *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

The presence of a unit root can be strongly rejected based on the Augmented Dickey-Fuller test for all series except hydrobalance. Interestingly, the KPSS test with the opposite null hypothesis of stationarity can also be rejected for power price and temperature series, although at the 10% significance level only. Haldrup and Nielsen (2006) report similar results in conducting a wide range of stationarity tests

on Nordic power prices, showing that neither $I(0)$ nor $I(1)$ processes seem to be appropriate. Let us get further insights into the time series properties of our data by examining Figure 4, which plots the sample autocorrelation functions (ACFs) for de-seasonalised power price, de-seasonalised temperature, and hydrobalance.

We can see slight peaks in power price ACF at lags 7 and 14, which indicate the presence of weekly seasonal effects not captured by Λ_1. Hydrobalance shows an extremely high degree of persistence in the autocorrelation function, typical for $I(1)$ processes. Both power price and temperature show a slow (hyperbolic) decay in autocorrelations that cannot be captured by traditional Autoregressive Moving Average (ARMA) models. This type of behaviour is known as long memory, or long-range dependence.

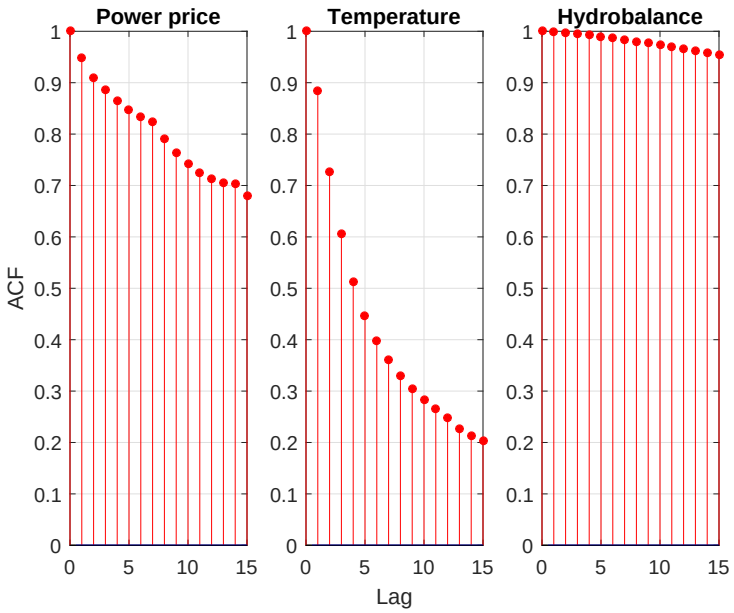


Figure 4: Sample autocorrelation functions.

Saying that a given process has long memory means that the effect of a single shock is extremely persistent. However, unlike in the case of a random walk, persistence can be combined with mean-reversion in long-memory models. There is no reason to assume that any of the series in question could wander arbitrarily away from

their fundamental mean levels, and therefore, a model that enforces mean-reversion while allowing high shock persistence appears to be a good choice.

Figure 5 illustrates the daily changes in de-seasonalised power log-price (i.e., log-returns), de-seasonalised temperature and hydrobalance. We can see pronounced clustering effects in power volatility, and seasonal patterns in temperature volatility with peaks during winter periods. Daily hydrobalance fluctuations have notable positive skewness, which suggests the need for a non-symmetric distribution. In addition, the graph reveals a frequent occurrence of extreme observations in power log-returns.

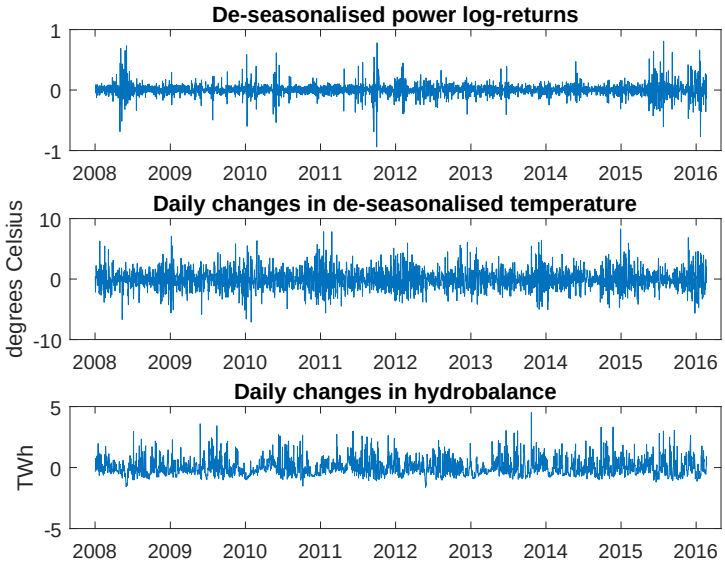


Figure 5: Time series of daily changes.

2.3. Relationship between power price, temperature, and hydrobalance

In the previous subsection, we investigated the properties of power prices, temperature, and hydrobalance separately. We now discuss how they are related, and the kind of dependence structure we want our model to impose.

Temperature is the main demand-side stochastic factor affecting power price. Temperature determines demand for heating during the winter periods and demand

for cooling during the summer periods. Ebbeler, Benth and Kiesel (2014) find that the correlation between de-seasonalised temperature and German power spot price is negative in the winter months, and positive, though lower in magnitude, in the summer months. Seasonal effects of this kind can be accommodated in the linear model, either by allowing the temperature coefficient in the power mean equation to take different values during cooling and heating seasons, or by having a single temperature coefficient for the entire sample and an additional coefficient for heating seasons only. We test both specifications by regressing de-seasonalised power log-returns on de-seasonalised temperature and find that there is no extra heating season effect. In fact, temperature has no effect on power evolution during summer months in our sample, so one could argue that the entire effect is driven by demand for heating. When considering October – March periods only, the estimated coefficient is negative and very close to the coefficient for the whole sample period. Thus, we conclude that Nordic summer temperatures are too mild to generate any significant cooling demand effect in the power market, and therefore, a single coefficient for each temperature lag in the power mean equation is sufficient.

Halldin (2005) discussed the inverse relationship between the Nordic power price and the hydro reservoir level in the context of the stochastic optimization of a hydro-thermal power system. We now compare the Nord Pool system prices under different hydrological conditions, but the same demand conditions. Here, we use the Nord Pool consumption data in addition to the data described previously. Figure 6 presents a scatterplot of power prices against power consumption, where we group observation pairs depending on the hydrobalance level. The red dots represent the lower quartile of hydrobalance, i.e., the driest 25% of days during the sample period. The blue crosses, on the other hand, mark the upper quartile of hydrobalance, i.e., the wettest 25% of days.

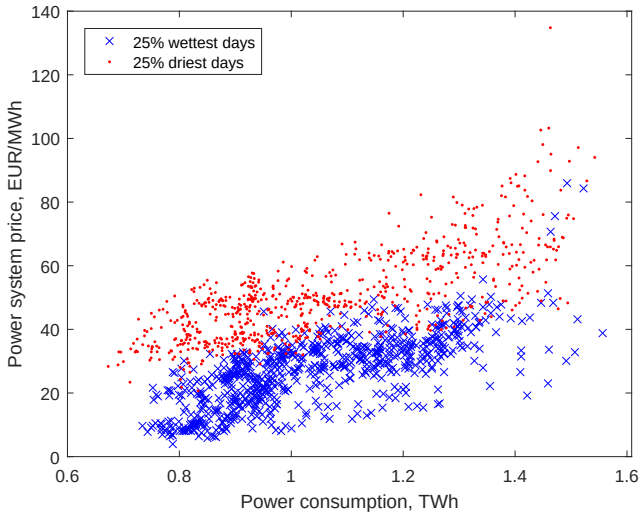


Figure 6: Power price versus power consumption scatterplot.

It is evident that given the same consumption, power price is generally higher under dry hydrological conditions. This is explained by the fundamentals of power price formation: as less hydro power is available, more generation technologies with higher marginal costs are utilised. In addition to a higher average price level, dry hydrological conditions lead to an increased probability of extreme prices. Additional regression analysis reveals that the relationship between power price and hydrobalance can be approximated sufficiently well by a single linear term, and there seems to be no motivation for including any non-linear effects.

To summarize, we would like the model to allow temperature and hydrobalance to influence the power price dynamics, but not the other way round. We expect a reasonable model to produce negative correlations between power and hydrobalance. Further, we expect the correlations between power and temperature to be negative during the heating seasons and close to zero otherwise.

3. The model

This section presents the model for the joint evolution of power price, temperature, and hydrobalance. We first describe the model in its general form, and motivate

how it is expected to capture the numerous features discovered in the preliminary analysis. Further, we discuss the model identification and estimation methodology.

3.1. General framework

3.1.1. Conditional mean

Considering all of the theoretical and empirical arguments outlined in Section 2, we suggest a model of the mean evolution of the system within the Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA) framework. Univariate ARFIMA processes are a well-known class of long-memory models, introduced by Granger and Joyeux (1980) and Hosking (1981) as a generalization of the traditional ARIMA (p, d, q) model, which allows the difference parameter d to take fractional values. Beran (1994) and Palma (2007) discuss the statistical properties and inference for long-memory processes in detail. Prior to becoming a subject of interest for econometricians, the fractional behaviour of certain time series was extensively studied in hydrology and climatology. The first published papers describing how to test and model the long-range dependence date back to the 1950s, and are concerned with modelling the inflows of the river Nile (see Hurst, 1951).[2] Brody et al. (2002) document fractional behaviour in the English temperature series. Haldrup and Nielsen (2006) explain that $I(d)$ processes with fractional d fit the power price data well in the presence of long memory effects. Similar to this study, Caporin et al. (2012) apply the VARFIMA framework to a joint model of power prices and temperatures.

We denote the power log-price series by $x_{1,t}$, the temperature series by $x_{2,t}$, and the hydrobalance series by $x_{3,t}$. Their joint evolution is governed by the following dynamic system:

$$\Phi(L)\Delta(L)\begin{pmatrix} x_{1,t} - \Lambda_{1,t} \\ x_{2,t} - \Lambda_{2,t} \\ x_{3,t} \end{pmatrix} = \Theta(L)\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}, \tag{3}$$

[2]Hurst was an English civil servant sent to Egypt as a hydrological consultant to predict how much the Nile floods from year to year. He developed rescaled range statistics, which became known later as the Hurst's exponent (H), and is related to the fractional difference parameter d.

$$\Phi(L) = I - \sum_{j=1}^{p} \begin{pmatrix} \phi_{11}^j & \phi_{12}^j & \phi_{13}^j \\ 0 & \phi_{22}^j & 0 \\ 0 & 0 & \phi_{33}^j \end{pmatrix} L^j, \ \Theta(L) = I + \sum_{k=1}^{q} \begin{pmatrix} \theta_{11}^k & \theta_{12}^k & \theta_{13}^k \\ 0 & \theta_{22}^k & 0 \\ 0 & 0 & \theta_{33}^k \end{pmatrix} L^k,$$

where $\Phi(L)$ is a restricted vector autoregressive (VAR) polynomial of order p, with L denoting the lag operator ($L^j x_t = x_{t-j}$), and I denoting a 3×3 identity matrix. $\Delta(L)$ is a diagonal long-memory matrix with a typical diagonal element $(1-L)^{d_i}$, d_i is a fractional difference parameter for variable i; $\Theta(L)$ is a restricted vector moving average (VMA) polynomial of order q; $(\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t})'$ is the vector of innovations; and finally, $\Lambda_{1,t}$ and $\Lambda_{2,t}$ are the seasonal mean functions given by Eq. (1) and Eq. (2), respectively. Recall that hydrobalance is de-seasonalised a priori, representing the deviations of the total hydrological resources from the seasonal mean level. $\Phi(L)$ and $\Theta(L)$ are restricted such that temperature and hydrobalance can affect power prices, while the opposite is not possible.[3] Note that in the process of lag order selection, we will restrict $\Phi(L)$ and $\Theta(L)$ even further to achieve the highest possible sparsity while retaining the essential effects. The parameters d_i determine the long-range behaviour of the series, while the parameters in $\Phi(L)$ and $\Theta(L)$, together with the lag order p and q, determine the short-range properties.

The difference operator $(1-L)^d$, for any real d, is an infinite linear filter given by the following binomial expansion:

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k L^k, \tag{4}$$

with the binomial coefficients $\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$, where $\Gamma(\cdot)$ denotes the Gamma function. Hosking (1981) shows that under certain assumptions ensuring stationarity and invertibility, a fractional process has infinite moving average and autoregressive representations with coefficients based on the binomial expansion of the difference operator. In practice, the truncated versions of these representations

[3] It is reasonable to allow temperature to influence the evolution of hydrobalance in the mean equation. However, since the primary focus of this study is power price dynamics, we do not explore the effect of temperature on hydrobalance.

are often estimated with approximate maximum likelihood methods. Not all fractionally integrated processes have long memory. An $I(d)$ process shows the long-memory property in the form of hyperbolic autocorrelation decay rate only if $d > 0$. Further, $d < \frac{1}{2}$ corresponds to a stationary process with a finite variance and an integrable spectral density. For the range $\frac{1}{2} \leq d < 1$, an $I(d)$ process is not stationary, but still mean reverting, and we can define its spectral density in a more general form, although not integrable (see Beran, 1994, for details). Finally, $d > 1$ leads to a non-stationary and non-mean-reverting case with long memory. Note that we can reduce the case of $d > 1$ to one of the cases mentioned above by taking the appropriate number of integer differences (e.g., if x_t is $I(1.2)$, then $(1 - L)^1 x_t = x_t - x_{t-1}$ is $I(0.2)$).

3.1.2. Conditional covariance

The innovation process in our model follows a conditional distribution with zero mean and time-varying covariance matrix H_t:

$$(\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t})' \mid \omega_{t-1} \sim D(0, H_t), \tag{5}$$

where ω_{t-1} denotes the information set at $t - 1$, which constitutes all past observations. The importance of modelling the time-varying volatility both in financial and meteorological data is an established fact. There is far less consensus, however, on what the best way to do this is, and the choice of model is often driven by the specifics of the dataset and the application in mind. Previous studies on temperature modelling suggest that temperature volatility has yearly cycles, similar to the mean. Benth and Šaltytė-Benth (2005, 2007) calibrate the truncated Fourier series to the daily temperature residuals. Campbell and Diebold (2005) propose conditional volatility dynamics for temperature that combines a seasonal component captured by Fourier series and a cyclical component captured by a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) process. Including a periodic component in the volatility process seems to be relevant for meteorological series. However, a very limited number of multivariate volatility models allow for inclusion of exogenous variables and deterministic terms, mostly due to the excessive parameter restrictions

required to guarantee positive semi-definiteness of the covariance matrix. In the setup most similar to ours, Caporin et al. (2012) model conditional variances of power and temperature by long-memory log-GARCH processes with deterministic components, and model conditional correlations separately. They mention that their approach to modelling correlations cannot be generalized to systems of a dimension higher than two.

In this study, we specify the evolution of the full covariance matrix directly using the multivariate GARCH framework of the Baba-Engle-Kraft-Kroner (BEKK) type defined in Engle and Kroner (1995). Following the multivariate GARCH literature, the vector of model error terms from Eq. (3) is written as:

$$\varepsilon_t = H_t^{1/2} z_t, \tag{6}$$

where z_t is a 3×1 vector of independently identically distributed (i.i.d.) innovations with zero mean and unit variance, and $H_t^{1/2}$ is the 3×3 square root of the conditional covariance matrix, which imposes the desired dependence structure.

We assume that the conditional covariance matrix H_t follows a BEKK(1,1) process:

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B, \tag{7}$$

where A, B, and C are 3×3 parameter matrices, C is lower triangular, and ε_{t-1} is the 3×1 vector of innovations in Eq. (3).

An important advantage of the BEKK model is that H_t is positive semi-definite by construction. In addition, the off-diagonal elements in the A and B matrices have immediate interpretations in terms of the cross-variable volatility spillover effects. Due to the nature of our series, we restrict some of these off-diagonal elements to zero. In particular, we rule out any cross-effects to the temperature series and allow temperature, but not power, to affect hydrobalance.

3.1.3. Distributional assumptions

We complete the model framework with a specification of the joint distribution of the i.i.d. innovation vector z_t in Eq. (6).

Despite the wide acknowledgement that financial data series exhibit heavy tails and skewness, the normal distribution is still dominant in the modelling literature for several reasons. First, it is convenient to resort to the asymptotic properties of the Quasi-Maximum Likelihood (QML) estimator, which is consistent even if the true conditional distribution of innovations is not normal, provided that the conditional mean and variance models are correctly specified. Second, the normal distribution often allows for closed-form pricing and hedging of derivative assets while introducing any non-normal dynamics requires computationally intensive numerical methods to price even standard derivatives in most cases.

Temperature series, on the other hand, can be much better approximated by the normal distribution than any price series. Most of the temperature modelling papers we referred to (Alaton, 2002; Brody et al., 2002; Campbell and Diebold, 2005) rely on the normal distribution assumption for the residuals. However, Benth and Šaltytė-Benth (2005) show that normality is rejected for some of the Norwegian temperature data and propose to apply the generalized hyperbolic distribution family. A closer look at Figure 2, which plots the Swedish temperature series, reveals that extreme deviations from the seasonal mean are quite common, especially in the winter periods. Recall that it is those extreme values, and not the average dynamics, that give rise to excess power demand and are of primary interest for any risk management or production planning purposes.

Finally, hydrological time series are known to be significantly positively skewed (see Helsel and Hirsch, 2002), which is also the case with our hydrobalance data, as confirmed by Figure 5. Considering all of the arguments above, we suggest using a flexible heavy-tailed and skewed distribution while keeping the normal distribution as the benchmark for comparison. The multivariate skew-Student distribution with independent components of Bauwens and Laurent (2002, 2005) appears to be an excellent choice, since it allows the univariate marginal distributions to have individual skewness and tail properties. Furthermore, this distribution is relatively straightforward to augment with GARCH-type dynamics.

In this study, we specify the multivariate skew-Student distribution for the vector of standardized innovations z_t in Eq. (6). Following Bauwens and Laurent (2002),

a $k \times 1$ random vector z_t is standard multivariate skew-Student distributed with independent components if its probability density function is given by:

$$f(z_t) = \left(\frac{2}{\sqrt{\pi}}\right)^k \left[\prod_{i=1}^{k} \frac{\xi_i s_i}{1+\xi_i^2} \frac{\Gamma\left(\frac{\upsilon_i+1}{2}\right)}{\Gamma\left(\frac{\upsilon_i}{2}\right)\sqrt{\upsilon_i-2}} \left(1+\frac{\kappa_{i,t}^2}{\upsilon_i-2}\right)^{-\frac{1+\upsilon_i}{2}}\right], \tag{8}$$

where

$$\kappa_{i,t} = (s_i z_{i,t} + m_i)\, \xi_i^{-I_{i,t}}, \tag{9}$$

and

$$I_{i,t} = \begin{cases} 1 \text{ if } z_{i,t} \geq -\frac{m_i}{s_i} \\ -1 \text{ if } z_{i,t} < -\frac{m_i}{s_i} \end{cases}, \tag{10}$$

with skewness parameters $\xi = (\xi_1, ..., \xi_k)$ and degrees of freedom parameters $\upsilon = (\upsilon_1, ..., \upsilon_k)$ for $\upsilon_i > 2$, and $\Gamma(\cdot)$ denoting the Gamma function. The constants $m_i = m_i(\xi_i, \upsilon_i)$ and $s_i = s_i(\xi_i, \upsilon_i)$ are the means and standard deviations of the non-standardized skew-Student density, respectively, defined by:

$$m_i(\xi_i, \upsilon_i) = \frac{\Gamma\left(\frac{\upsilon_i-1}{2}\right)\sqrt{\upsilon_i-2}}{\sqrt{\pi}\Gamma\left(\frac{\upsilon_i}{2}\right)}\left(\xi_i - \frac{1}{\xi_i}\right), \tag{11}$$

$$s_i^2(\xi_i, \upsilon_i) = \left(\xi_i^2 + \frac{1}{\xi_i^2} - 1\right) - m_i^2. \tag{12}$$

$\xi_i = 1$ corresponds to the symmetric density, while the thickness of the tails is decreasing in υ_i. Note that the standardized multivariate normal density is the limiting distribution of $f(z_t)$ in Eq. (8), when $\xi_i = 1$ and $\upsilon_i \to \infty$.

3.2. Model identification and estimation procedure

Long-memory model estimation is a well-addressed area, and many estimation methods have been proposed in the literature. Most are based on either a time domain or frequency domain representation of the density function. The time domain procedures include various implementations of exact maximum likelihood, such as the Durbin-Levinson algorithm and state space methods, as well as a number of approximate likelihood methods based on truncated versions of autoregressive and

moving average representations of long-memory processes (see, e.g., Hasslett and Raftery, 1989). The frequency domain procedures include Whittle estimators and various semiparametric methods, and are based in one way or another on the calculation of the periodogram of the series using Fast Fourier Transform (FFT). These methods offer significant computational advantage over the time domain methods, but at the cost of lower precision of estimates. In this study, we estimate all model parameters jointly in the time domain using exact maximum likelihood. However, model complexity requires us to proceed in several steps.

First, we estimate the parameters in Λ_1 and Λ_2 using the least squares method. Subtracting these functions from the original power log-price and temperature observations yields the de-seasonalised series, which are the input to the next step.

The second step is to get the initial estimates of the fractional difference parameters. We would like to get a consistent estimate of the degree of fractional integration in the series without making any prior assumptions about the short-range properties. This can be achieved by using a semiparametric estimation method, which does not require specification of the parametric model and relies only on the assumption about the shape of the spectral density of the time series. The most common semiparametric methods to estimate long-memory parameters are local Whittle (see Künsch, 1987, and Robinson, 1995a), and log-periodogram regression (see Geweke and Porter-Hudak, 1983, and Robinson, 1995b). However, as Shimotsu and Phillips (2005) point out, these estimators are inconsistent for $d > 1$, and discontinuous at several points in the non-stationary region, leading to non-normal limit theory. Instead, they suggest a general purpose semiparametric estimator called the exact local Whittle estimator with well-behaved asymptotic properties in the wide range of stationary and non-stationary values. We use the exact local Whittle estimator of Shimotsu and Phillips (2005) to obtain the initial estimates of the d-parameters.

The third step is to apply the fractional difference filter to the series and identify the short-range part of the model, that is, the structure of $\Phi(L)$ and $\Theta(L)$. Calculating fractional differences is in itself a non-trivial task. Standard implementations of fractional differencing based on the binomial expansion of the difference operator have $\mathcal{O}(n^2)$ time complexity, which means that the number of operations performed

to compute the differenced series is a quadratic function of the input size. This is acceptable if differences are only to be computed once, but makes the joint estimation of all parameters in a trivariate model with several thousand time series points practically infeasible. However, Jensen and Nielsen (2013) suggest a fast fractional difference algorithm that takes advantage of a frequency-domain transform of the series. Their algorithm is of $\mathcal{O}(n \log n)$ time complexity and offers substantial computational advantages. We identify the short-range dynamics by inspecting the ACFs and PACFs of the differenced series following the standard practice. In addition, we conduct a number of univariate estimations assuming constant variance and compare them based on the information criteria.

The final step is the joint estimation of all model parameters by exact maximum likelihood using the parameter estimates from the previous steps as starting values only.[4] We implement this procedure for the cases of normal distribution and skew-Student distribution, separately. In the case of normally distributed residuals, the log-likelihood function is given by the log of the multivariate normal density function. In the case of the skew-Student distributed residuals, the log-likelihood function is given by:

$$\ln L(\theta) = \sum_{t=\max(p,q)+1}^{T} \left\{ \ln f(z_t) - \frac{1}{2} \ln |H_t| \right\}, \tag{13}$$

where θ is the parameter vector for the full model, $f(z_t)$ is the probability density function in Eq. (8), T is the number of time series observations, and $|H_t|$ denotes the determinant of H_t. Note that the summation is conditional on the first p or q observations, whichever is larger, owing to the lag order of $\Phi(L)$ and $\Theta(L)$ in the mean equations. The second term in the sum in Eq. (13) is the Jacobian correction term arising in the transformation from z to ε. We calculate the square root matrix $H_t^{1/2}$, which is required to obtain the vector of standardized residuals z_t as given by Eq. (6) at each time point using a standard spectral decomposition. The initial H_t

[4] Parameters in the seasonal mean functions are not re-estimated to decrease computational time. As starting values for the shape parameters in the skew-Student distributions, we use $\xi_i = 1$ for all skewness parameters and $\upsilon_i = 100$ for all kurtosis parameters ($\upsilon_i \to \infty$ corresponds to normality).

is set to the sample covariance matrix of the fractionally differenced data and the initial values of the residuals are set to zero.

The log-likelihood function is maximized by simulated annealing, following Goffe, Ferrier and Rogers (1994). The fundamental property of simulated annealing is that it is allowed to accept worse intermediate solutions (downhill moves) while searching for the optimum, which leads to a more extensive exploration of the parameter space and prevents the algorithm from becoming stuck in local optima. In theory, this property also makes the algorithm insensitive to starting values. To further increase the chance of identifying the global optimum, we use consistent QML estimates as starting values for the model with skew-Student distributed innovations. Finally, we calculate the standard errors of the parameters using the outer product gradient method with numerical first derivatives.

4. Results

This section presents the results of model identification and estimation. We start by discussing how we identified the conditional mean system within our general framework. We then discuss the estimated parameters. Finally, we examine the model implied second moments in light of the findings from our preliminary analysis.

4.1. Model identification results

The long-memory parameters are estimated by the exact local Whittle method as follows: 0.6747 for power, 0.2970 for temperature, and 1.1115 for hydrobalance. These results are consistent with the autocorrelation functions of the non-differenced data in Figure 4. Specifically, hydrobalance has the slowest ACF decay, reflected in the highest d-parameter, while temperature shows the fastest (yet still hyperbolic) ACF decay, with the lowest d-parameter. Further, the d-parameter of temperature lies in the stationary region, while the other two take values in the non-stationary region. It is worth noting that the d-parameter of hydrobalance is above one, which means that mean the reversion property is lost. The implications of this result on the model simulation will be discussed further.

We proceed to identifying the short-range properties of the conditional mean system. Figure 7 displays the autocorrelation and partial autocorrelation functions of the fractionally differenced series. The 95% white noise confidence bounds are given by the horizontal blue lines.

We can see that power has a slight spike in both functions at the first lag, and weekly periodic patterns. One alternative to capture weekly periodicities is to take seasonal differences. However, due to the presence of non-seasonal fractional differencing in our model, we prefer to include seasonal autoregressive lags instead. Further investigation in the univariate framework reveals that two weekly seasonal terms, in addition to the non-seasonal AR(1) term, are sufficient to whiten the residuals. Moreover, a parsimonious specification with a single temperature term and a single hydrobalance term in the power mean equation is preferred based on the information criteria.

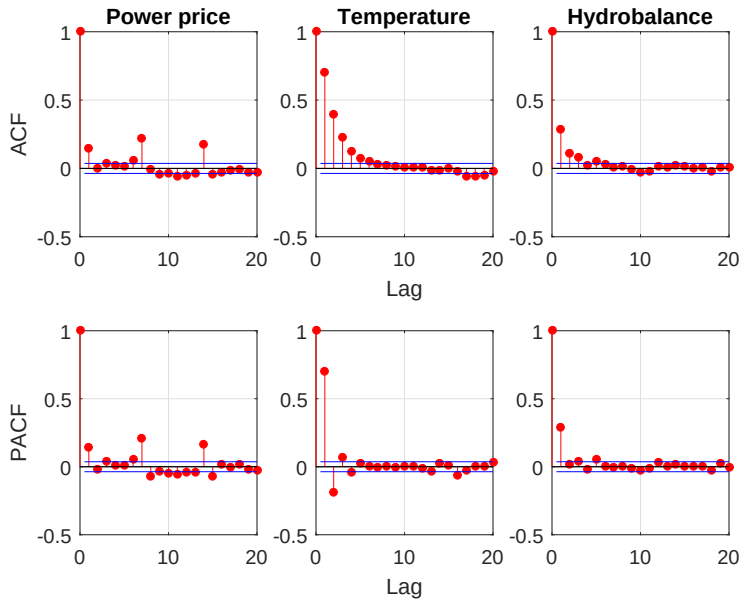


Figure 7: Autocorrelation and partial autocorrelation functions after fractional differencing.

Both temperature and hydrobalance show an AR signature, with ACF decaying gradually and PACF truncated at a certain lag. The PACF of temperature features

three significant lags, while the PACF of hydrobalance is truncated at the first lag. Overall, the interpretation of the functions after appropriate fractional differencing is straightforward, suggesting to model temperature as an AR(3) process and hydrobalance as an AR(1) process. We argue that there is no need to include any moving average terms, thus reducing the general VARFIMA framework presented in the previous section to its special case of fractionally integrated vector autoregression. This leads us to identify the following conditional mean equation system:

$$\begin{aligned} &y_{1,t} - \phi_{11}^{1} y_{1,t-1} - \phi_{11}^{7} y_{1,t-7} - \phi_{11}^{14} y_{1,t-14} - \phi_{12}^{1} y_{2,t-1} - \phi_{13}^{1} y_{3,t-1} = \varepsilon_{1,t} \\ &y_{2,t} - \phi_{22}^{1} y_{2,t-1} - \phi_{22}^{2} y_{2,t-2} - \phi_{22}^{3} y_{2,t-3} = \varepsilon_{2,t} \\ &y_{3,t} - \phi_{33}^{1} y_{3,t-1} = \varepsilon_{3,t} \end{aligned} \tag{14}$$

where y_1, y_2, and y_3 denote de-seasonalised fractionally differenced power log-price, de-seasonalised fractionally differenced temperature, and fractionally differenced hydrobalance, respectively.

4.2. Estimation results

We first comment on the seasonal mean parameter estimates presented in Table 2.

Table 2: Seasonal mean parameter estimates.

Power		Temperature	
$\lambda_{1,1}$: constant	3.9575	$\lambda_{2,1}$: constant	6.9372
	(0.0113)		(0.0614)
$\lambda_{1,2}$: amplitude	0.1606	$\lambda_{2,2}$: amplitude	−9.7656
	(0.0085)		(0.0861)
$\lambda_{1,3}$: phase angle	7.9960	$\lambda_{2,3}$: phase angle	21.728
	(3.1498)		(0.5202)
$\lambda_{1,4}$: linear trend	−0.0003		
	(0.0000)		
$\lambda_{1,5}$: non-business days	−0.1205		
	(0.0132)		

Note: The table reports the estimated coefficients and their standard errors (in parentheses). The seasonal mean functions are given by Eq. (1) and Eq. (2). All coefficients are statistically significant at the 1% level. The R^2 values are 0.400 for power and 0.846 for temperature.

We can interpret the parameters $\lambda_{1,1}$ and $\lambda_{2,1}$ as constant average levels of power log-price and temperature, respectively. The estimated average log-price corresponds to the average price of 52.32 EUR/MWh. The parameters $\lambda_{1,2}$ and $\lambda_{2,2}$ represent half of the distance between the peaks and the troughs of the yearly seasonal functions. Thus, the difference between the temperature highs and lows is around 19 degrees Celsius, while the average power price difference between warm and cold seasons is 16.88 EUR/MWh.[5] Parameters $\lambda_{1,3}$ and $\lambda_{2,3}$ act as time shifts, placing the cosine waves in the correct phase of power price and temperature yearly cycles. Further, we report a significant decreasing linear time trend in the power price series. As mentioned in Section 2, this phenomenon is related to the changes in the Nordic power generation mix, with a growing share of load covered by low marginal cost renewable generation sources. Finally, note the non-trivial effect of non-business days, comparable in magnitude to the amplitude of the yearly cycle.

We estimated the remaining parameters under two distributional assumptions: the normal distribution, serving as the benchmark case, and the more flexible skew-Student distribution. Since the two model specifications are not nested, it is not possible to formally test them against each other using the likelihood ratio test. However, we can still get an idea of the gain from departing from normality by comparing the starting log-likelihood value of the skew-Student specification, which is -4559.23, with the final log-likelihood value of -3696.94. Recall that the optimal parameter values from the normal specification were the starting values for the skew-Student specification, while the starting values of the ξ- and υ-parameters were set to roughly correspond to normality.

Table 3 presents the estimated values of the parameters from the stochastic component of the conditional mean system.

The estimates of the memory parameters are very close to the univariate exact local Whittle estimates discussed earlier. Haldrup and Nielsen (2006) find that the Nordic zonal hourly spot price series show long memory with d ranging between 0.31 and 0.52. Caporin et al. (2012) report d-parameters of 0.39 and 0.19 for the Oslo

[5] $e^{(3.9575+0.1606)} - e^{(3.9575-0.1606)}$

area daily power and temperature series. We find a higher degree of long memory for both series, with d_1 of power being well in the non-stationary region.

Table 3: Stochastic conditional mean parameter estimates.

Normal						Skew-Student					
Power		Temperature		Hydrobalance		Power		Temperature		Hydrobalance	
d_1	0.6963	d_2	0.2660	d_3	1.2103	d_1	0.6445	d_2	0.2957	d_3	1.1478
	(0.0231)		(0.0371)		(0.0227)		(0.0214)		(0.0360)		(0.0145)
ϕ_{11}^1	*0.0130*	ϕ_{22}^1	0.8506	ϕ_{33}^1	0.1351	ϕ_{11}^1	0.0755	ϕ_{22}^1	0.8283	ϕ_{33}^1	0.1596
	(0.0290)		(0.0411)		(0.0280)		(0.0278)		(0.0394)		(0.019)
ϕ_{11}^7	0.2128	ϕ_{22}^2	-0.2561			ϕ_{11}^7	0.2058	ϕ_{22}^2	-0.2367		
	(0.0155)		(0.0259)				(0.0143)		(0.0244)		
ϕ_{11}^{14}	0.1289	ϕ_{22}^3	0.0629			ϕ_{11}^{14}	0.1170	ϕ_{22}^3	0.0445		
	(0.0124)		(0.0193)				(0.0120)		(0.0183)		
ϕ_{12}^1	-0.0037					ϕ_{12}^1	-0.0035				
	(0.0005)						(0.0005)				
ϕ_{13}^1	-0.0153					ϕ_{13}^1	-0.0126				
	(0.0014)						(0.0015)				

Note: The table reports the estimated coefficients and their standard errors (in parentheses). The conditional mean system is given by Eq. (3) in the general form for the original series, and Eq. (14) in the restricted form for the deseasonalised and differenced series. Non-significant coefficients are reported in italics, while the remaining coefficients are statistically significant at the 1% level.

In general, there are no extreme differences between the parameter estimates from the two distributional specifications. We can see the natural trade-off between the degree of long memory and the magnitude of the non-seasonal autoregressive coefficients in all the three series. In particular, a higher estimate of d_1 under the normal distribution is coupled with a lower and nonsignificant ϕ_{11}^1. The autoregressive coefficients from the temperature mean equation are slightly lower in magnitude than those reported in Benth, Šaltytė-Benth and Koekebakker (2008) for the Stockholm temperature series, due to the presence of the long-memory component in our specification. Further, we report significant ϕ_{11}^7 and ϕ_{11}^{14} coefficients capturing the first- and the second-order weekly seasonal autoregressive patterns in the power price. Finally, we find significant negative temperature and hydrobalance effects in the power mean equation, which is in line with our preliminary analysis.

Table 4 presents the estimated values of the conditional covariance and distributional parameters.

We first analyse the diagonal coefficients of A and B-matrices. Caporin et al. (2012) find that temperature has a lower degree of persistence in volatility than power price. According to our results, temperature and hydrobalance volatility have weaker ARCH-effects (as measured by a_{22} and a_{33}) compared to power price volatility. However, the GARCH-coefficient b_{22} for temperature suggests a higher degree of persistence in volatility than the corresponding coefficient b_{11} for power. This might be related to the fact that we do not explicitly model the seasonality in temperature volatility, so part of it is accommodated by b_{22}.

Table 4: Conditional covariance and distributional parameter estimates.

Normal						Skew-Student					
c_{11}	0.0102					c_{11}	0.0153^{***}				
	(0.0134)						(0.0017)				
c_{21}	-0.0189^{***}	c_{22}	0.1169^{***}			c_{21}	-0.0069^{***}	c_{22}	0.1904^{***}		
	(0.0072)		(0.0377)				(0.0020)		(0.0274)		
c_{31}	-0.0076^{***}	c_{32}	-0.1312^{***}	c_{33}	0.0791^{***}	c_{31}	-0.0088^{***}	c_{32}	-0.0792^{***}	c_{33}	0.0752^{***}
	(0.0020)		(0.0244)		(0.0044)		(0.0016)		(0.0221)		(0.0084)
a_{11}	0.5435^{***}	a_{12}	-0.5599^{**}	a_{13}	-0.1963^{**}	a_{11}	0.4414^{***}	a_{12}	-0.1430	a_{13}	-0.2178^{***}
	(0.0142)		(0.2255)		(0.0956)		(0.0223)		(0.1733)		(0.0691)
		a_{22}	0.1830^{***}					a_{22}	0.1615^{***}		
			(0.0117)						(0.0137)		
		a_{32}	-0.0609^{**}	a_{33}	0.1350^{***}			a_{32}	-0.0196	a_{33}	0.1956^{***}
			(0.0274)		(0.0067)				(0.0258)		(0.0126)
b_{11}	0.8311^{***}	b_{12}	0.2075^{*}	b_{13}	0.0561	b_{11}	0.8786^{***}	b_{12}	-0.0272	b_{13}	0.1554^{***}
	(0.0086)		(0.1071)		(0.0400)		(0.0094)		(0.0707)		(0.0288)
		b_{22}	0.9734^{***}					b_{22}	0.9764^{***}		
			(0.0032)						(0.0038)		
		b_{32}	0.0471^{***}	b_{33}	0.9829^{***}			b_{32}	0.0224^{***}	b_{33}	0.9764^{***}
			(0.0078)		(0.0016)				(0.0074)		(0.0028)
						ξ_1	0.9843^{***}	ξ_2	1.0288^{***}	ξ_3	1.8734^{***}
							(0.0212)		(0.0270)		(0.0574)
						υ_1^{-1}	0.2598^{***}	υ_2^{-1}	0.1351^{***}	υ_3^{-1}	0.2569^{***}
							(0.0168)		(0.0218)		(0.0198)

Note: The table reports the estimated coefficients and their standard errors (in parentheses). The conditional covariance model is given by Eq. (7) and is parameterized by the 3×3 matrices C, A and B, with typical elements c_{ij}, a_{ij}, and b_{ij}, for $i, j = 1$ (power), 2 (temperature), and 3 (hydrobalance), respectively. The C-matrix is lower triangular, while the A and B matrices are restricted such that temperature volatility dynamics is exogenous, hydrobalance volatility is allowed to be affected by temperature but not by power, while all variables in the system can affect power volatility. Skew-Student distributional parameters are reported in the last two rows. Superscripts *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

The majority of the off-diagonal coefficients both in A and B-matrices are statistically significant, which confirms the existence of volatility spillover effects. There are differences, however, in the estimates of these effects between our two distributional specifications. The normal specification features significant spillovers from temperature to power volatility, as measured by a_{12} and b_{12}, and smaller magnitude a_{13} coefficient, representing the spillover from hydrobalance to power volatility. The skew-Student specification, on the other hand, results in nonsignificant temperature-to-power effects, but highly statistically significant and sizeable hydrobalance-to-power effects, as measured by both a_{13} and b_{13}.

We also find the b_{32}-coefficient on temperature-to-hydrobalance volatility spillover effect to be statistically significant at the 1% level in both specifications. However, its magnitude is much lower than of the coefficients capturing spillovers to power. Note that the signs of the off-diagonal parameters do not have a straightforward interpretation because these parameters appear in several non-linear terms determining each element of the conditional covariance matrix at each time point.

The last two rows in Table 4 report the estimates of the skewness and the inverses of the degrees of freedom parameters.[6] Although the ξ_1 of power is below 1, and the ξ_2 of temperature is above 1, the 95% confidence intervals for these parameters leave the question of asymmetry open, with the lower bound in the negative region and the upper bound in the positive region. The story is different with the ξ_3 of hydrobalance, which is well in the positive asymmetry region, as expected. Further, we find that power and hydrobalance have very similar tail properties with υ-parameters close to 4. Temperature shows less heavy tails with the υ_2 estimate of 7.4, although it still implies a relatively fat-tailed distribution. Taking another look at Figure 2, we expect that this estimate is mostly driven by the extreme temperature occurrences in winter periods, and this has to be taken into account in a simulation from the model.

Diagnostic checks of the residuals from both model specifications reveal that we are left with zero-mean uncorrelated noise. Examining the estimated volatilities and

[6] We estimate υ^{-1} instead of υ itself for numerical reasons.

correlations, plotted in Figures 8 and 9, allows us to further assess the in-sample model performance.

We can see the resemblance between the estimated volatility processes and the daily changes series in Figure 5. Temperature volatility displays peaks during the winter periods and troughs during the summer periods. Hydrobalance volatility starts the yearly cycle at a relatively low level, reaches the minimum around April, and then takes on an upward trend with a peak in August – September, followed by a sharp decline.

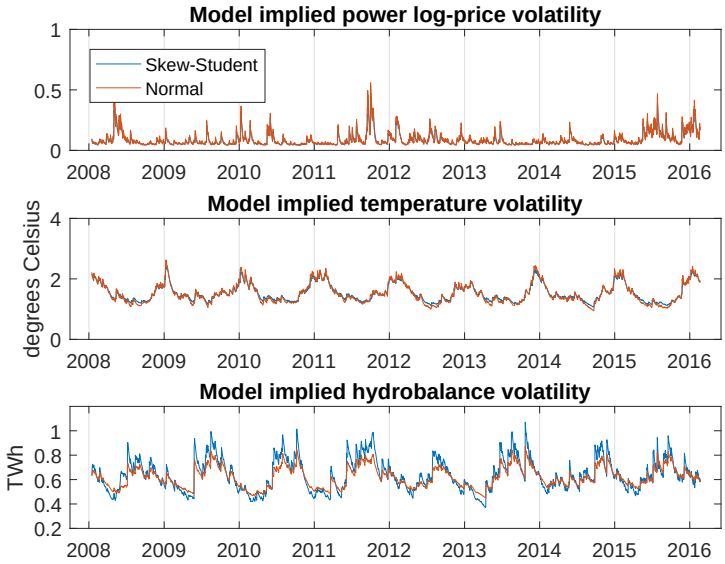


Figure 8: Model implied volatilities.

The estimated conditional correlations are well in line with our expectations. The correlation between power and temperature reaches the minimum of -0.5 to -0.4 during the winter months, and is roughly zero during the summer months. The correlation between power and hydrobalance does not have a pronounced seasonal shape, and mostly stays in the negative region between -0.5 and 0, oscillating around the average level of -0.25, with a few extremes.

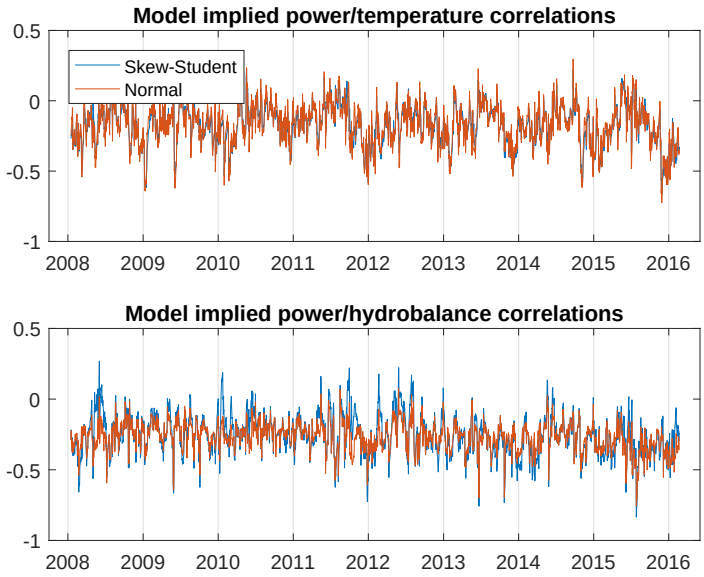


Figure 9: Model implied correlations.

Finally, there seem to be very minor differences between the estimated second moments from the two distributional specifications. Hydrobalance volatility and the power/hydrobalance correlations display more variation under the skew-Student specification. On the rest of the plots, the red lines and the blue lines coincide almost perfectly.

5. Application: Scenario analysis

This section illustrates how our model can be used to generate a number of power price scenarios under different hydrological and temperature conditions. We address the issues related to simulation from the model and present an overview of the simulation results from the skew-Student specification.

We start the simulation on February 22, 2016 (the day after the sample period ends), and finish on February 28, 2017, yielding a simulation length of 373 days. We use the last sample values of power price, temperature, and hydrobalance as the starting points for all simulated paths. In addition, the estimated conditional

covariance matrix on the last day of the sample period (H_T) and the last values of the residuals are used to iterate the BEKK process forward. For each day of the simulation period, we draw the random shocks z_t from the underlying skew-Student distributions using the analytical quantile function, given in Laurent (2002). Appendix A presents the kernel density estimates of the random samples drawn from the univariate skew-Student distributions with skewness and degrees of freedom parameters equal to our estimates. We construct the error terms ε_t using Eq. (6). Next, the long memory is created by applying numerical fractional integration of the error term series.[7] Further, we follow Eq. (14) to generate the stochastic mean component. Finally, we add the predicted seasonal mean component for power and temperature, and transform the log-price back to the natural units.

Figures 10 – 12 show the historical data series starting from April 27, 2015, followed by ten simulated paths.

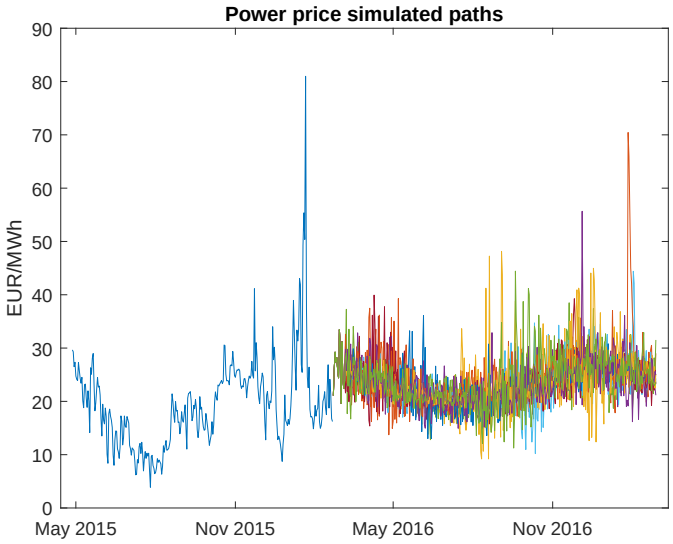


Figure 10: Historical power prices and simulated paths.

[7] To generate a fractionally integrated process, we approximate the binomial expansion of $(1 - L)^{-d}$ by truncating at 100 terms. The first 100 simulated values use the actual model residuals.

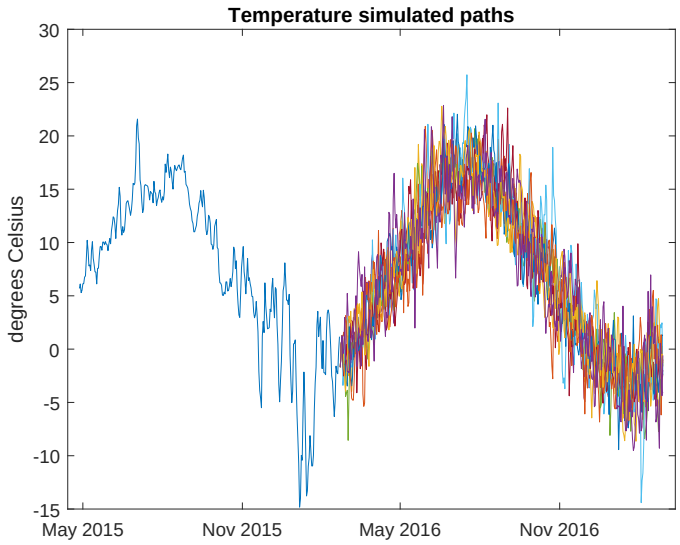


Figure 11: Historical temperature and simulated paths.

A first glance at the simulated paths suggests that the model does a fairly good job of capturing the empirical properties of the modelled series. It is worth noting that since the extreme power price and temperature observations are typically observed during the winter months, we can adjust the random shock generation to reflect this.

We can see in Figure 11 that only one temperature path out of ten produced an outlier during the winter period. However, we can easily adjust the simulation procedure to increase the likelihood of outliers occurring in winter, and if necessary, decrease the likelihood of outliers in other periods.[8] Overall, the simulation stage reveals the true benefits of using the skew-Student distribution, since the normal distribution cannot generate large enough moves frequently enough to produce realistic behaviour.

[8]For instance, this can be done by drawing the quantiles from a distribution other than uniform while generating the random values from the skew-Student distribution.

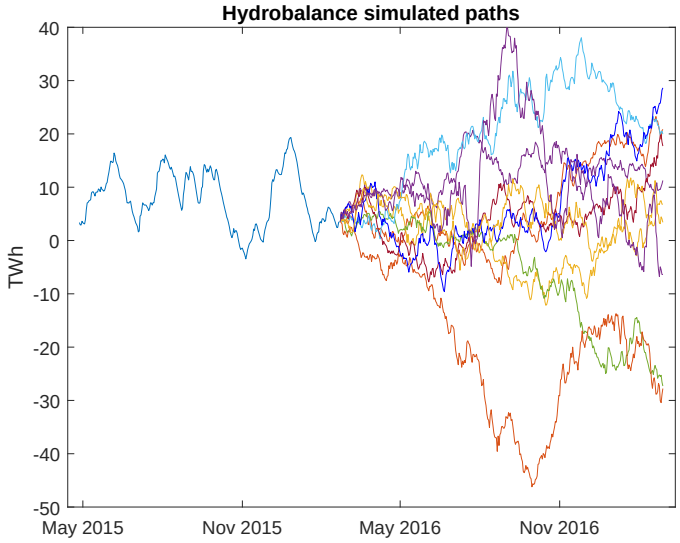


Figure 12: Historical hydrobalance and simulated paths.

As Figure 12 illustrates, we can generate a wide variety of hydrological scenarios. Recall from the discussion of the difference parameter properties that a value of d above 1 corresponds to a non-stationary case without mean reversion property. This means that some of the hydrobalance paths produced by our model will be unreasonable, and have to be discarded, which can also be automated in the simulation by introducing the bounds on how far hydrobalance can wander away from zero.

A limitation of the model that is more fundamental is that although extreme power price occurrences are likely, there is nothing in the model enforcing fast mean reversion and enabling spikes to appear once in a while. On the contrary, the otherwise desirable long-memory property makes it impossible for a single extreme value to occur. However, for an application such as meteorological scenario analysis, this limitation is of minor importance.

Our model creates plenty of interesting opportunities in scenario analysis. For instance, an average power price can be calculated for a range of temperature and hydrobalance combinations. This might be of interest for production planning in power markets with heavy reliance on hydrological generation, such as the Nordic

market. As an illustration, we simulate 1000 scenarios from the model and calculate the average power price, temperature, and hydrobalance values for November 1, 2016 – February 28, 2017, for each scenario. Figure 13 plots the average power price and temperature observation pairs grouped by hydrological conditions. The red dots mark the driest quarter of scenarios, while the blue crosses mark the wettest quarter. We omit the observations in the middle of the range.

We can see that, with rare exceptions, the average power price is higher in dry scenarios, given the same average temperature. Further, extremely high average power prices tend to occur under a combination of cold and dry conditions. A similar analysis can be done for statistics other than the average.

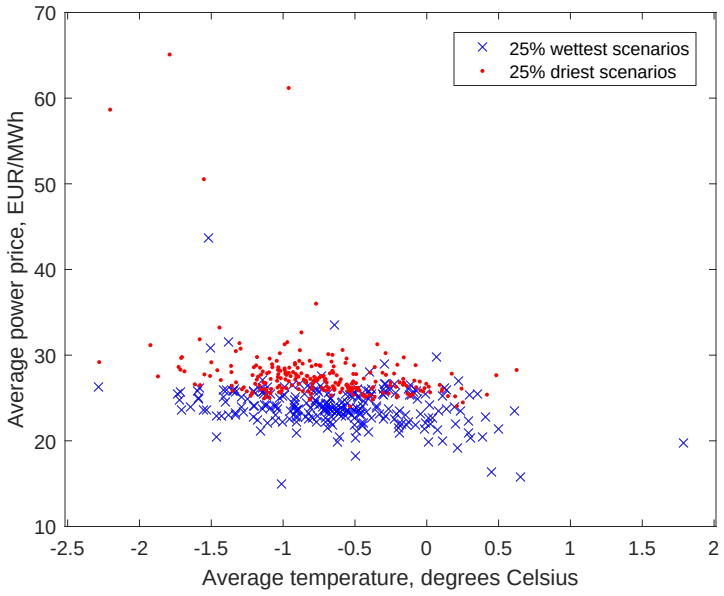


Figure 13: A scenario analysis example.

Another potential application of the model is Monte Carlo pricing of weather derivatives, such as energy quanto contracts. Energy quanto contracts have a payoff that depends on the product of two indices: an energy index (e.g., an average power price during the delivery period) and a temperature index (e.g., Heating Degree Days, or HDD). A wide variety of payoff structures is possible for quanto contracts, such as

swap, put/call, collar, and so on. Caporin et al. (2012) illustrate how their bivariate power-temperature model can be utilised to price such contracts. Our model can be applied in a similar fashion, with the additional flexibility of computing 'hydrological bounds' on contract prices.

6. Summary and conclusions

In this study we propose a model for the joint evolution of spot power price, temperature, and hydrobalance. Our model successfully captures most of the discovered empirical features, such as long memory and heavy tails in all series, yearly seasonal patterns in power price and temperature, weekly periodic patterns in power price, pronounced positive skewness in hydrobalance, and time-varying conditional second moments. We find that in the Nordic market, power price is inversely related to temperature throughout the year, except for summer months, when the effect is nonsignificant. Hydrobalance, on the other hand, negatively affects power price in all periods, since in dry hydrological conditions, higher marginal cost generation sources set the price. Further, we confirm the existence of volatility spillover effects from temperature and hydrobalance to power. We illustrate how our model can be used to generate a variety of weather scenarios and to analyse the implications for power prices. The model is relevant for power markets with a dominant share of hydrological generation and provides a wide scope of opportunities for scenario analysis with relatively little meteorological input.

References

[1] Alaton, P., Djehiche, B., Stillberger, D., 2002. On modelling and pricing weather derivatives. Applied Mathematical Finance 9, 1–20.

[2] Bauwens, L., Laurent, S., 2002. A new class of multivariate skew densities, with applications to GARCH models. CORE Discussion Paper 20.

[3] Bauwens, L., Laurent, S., 2005. A new class of multivariate skew densities, with application to Generalized Autoregressive Conditional Heteroscedasticity models. Journal of Business and Economic Statistics 23, 346–354.

[4] Benth, F.E., Šaltytė-Benth, J., 2005. Stochastic modelling of temperature variations with a view towards weather derivatives. Applied Mathematical Finance 12, 53–85.

[5] Benth, F.E., Šaltytė-Benth, J., 2007. The volatility of temperature and pricing of weather derivatives. Quantitative Finance 7, 553–561.

[6] Benth, F.E., Šaltytė-Benth, J., Koekebakker, S., 2008. Stochastic modelling of electricity and related markets. Advanced series on statistical science and applied probability; v. 11. World Scientific, Singapore.

[7] Benth, F.E., Lange, N., Myklebust, T.Å., 2012. Pricing and hedging quanto options in energy markets. SSRN Working Paper.

[8] Beran, J., 1994. Statistics for long-memory processes. Chapman and Hall, London.

[9] Brody, D.C., Syroka, J., Zervos, M., 2002. Dynamical pricing of weather derivatives. Quantitative Finance 3, 189–198.

[10] Campbell, S.D., Diebold, F.X., 2005. Weather forecasting for weather derivatives. Journal of the American Statistical Association 100, 6–16.

[11] Caporin, M., Pres, J., Torro, H., 2012. Model based Monte Carlo pricing of energy and temperature Quanto options. Energy Economics 34, 1700–1712.

[12] Ebbeler, S., Benth, F.E., Kiesel, R., 2014. Indifference pricing of weather derivatives based on electricity futures. In: Prokopczuk, M. (Ed.). Energy pricing models: recent advances, methods, and tools. Palgrave Macmillan, New York.

[13] Engle, R., Kroner, K., 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11, 122–150.

[14] Geweke, J., Porter-Hudak, H., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4, 221–238.

[15] Goffe, W., Ferrier, G., Rogers, J., 1994. Global optimization of statistical functions with simulated annealing. Journal of Econometrics 60, 65–99.

[16] Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1, 15–29.

[17] Green, R., 2015. A power market forward curve with hydrology dependence - an approach based on Artificial Neural Networks. Knut Wicksell Working Paper Series 2015/1, Knut Wicksell Centre for Financial Studies, Lund University.

[18] Haldrup, N., Nielsen, M.Ø., 2006. A regime switching long memory model for electricity prices. Journal of Econometrics 135, 349–376.

[19] Halldin, R., 2005. Stochastic modeling and optimization under uncertainty of a hydro power system. Doctoral dissertation, Centre for Mathematical Sciences, Lund Institute of Technology.

[20] Hasslett, J., Raftery, A.E., 1989. Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. Journal of Applied Statistics 38, 1–50.

[21] Helsel, D.R., Hirsch, R.M., 2002. Statistical methods in water resources. Techniques of water resources investigations, Book 4, chapter A3. U.S. Geological Survey.

[22] Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165–176.

[23] Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers 116, 770–799.

[24] Jensen, A.N., Nielsen, M.Ø., 2013. A fast fractional difference algorithm. QED Working Paper 1307, Queen's University.

[25] Künsch, H., 1987. Statistical aspects of self-similar processes. In: Prokhorov, Y., Sazanov, V. (Ed.): Proc. First World Congress of the Bernoulli Society 1, 67–74. VNU Science Press, Utrecht.

[26] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54, 159–178.

[27] Laurent, S., 2002. Asymmetry and fat-tails in financial time series. Doctoral dissertation, Maastricht University.

[28] Lucia, J., Schwartz, E., 2002. Electricity prices and power derivatives: evidence from the Nordic power exchange. Review of Derivatives Research 5, 5–50.

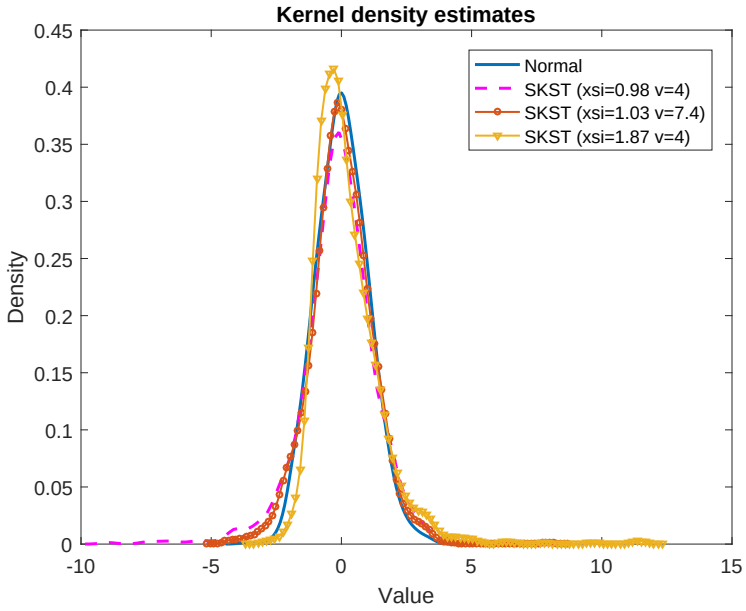
[29] Palma, W., 2007. Long-memory time series models: theory and methods. Wiley Series in Probability and Statistics. John Wiley and Sons, New Jersey.

[30] Robinson, P.M., 1995a. Gaussian semiparametric estimation of long range dependence. Annals of Statistics 23, 1630–1661.

[31] Robinson, P.M., 1995b. Log-periodogram regression of time series with long-range dependence. Annals of Statistics 23, 1048–1072.

[32] Shimotsu, K., Phillips, P.C.B., 2005. Exact local Whittle estimation of fractional integration. Annals of Statistics 33, 1890–1933.

Appendix A



Note: The figure shows the smooth kernel density estimates of the random samples drawn from the standard normal distribution and the standardized skew-Student distribution with ξ and υ parameters equal to our estimates. Each random sample consists of 1000 values.