

Faculty of Technology
Department of Mathematics and Physics
Laboratory of Applied Mathematics

Applying fluid mechanics and Kalman filtering to forecasting electricity spot prices

Anna Shcherbacheva

The topic of this Master's thesis was approved
by the faculty council of the Faculty of Technology on May 26, 2011.

The examiners of the thesis were: prof. Ph.D. Tuomo Kauranne and prof.
Ph.D. Heikki Haario

The thesis was supervised by prof. Ph.D. Tuomo Kauranne

Lappeenranta, August 11, 2011

Anna Shcherbacheva
Ruskonlahdenkatu 13-15, F19
53850 Lappeenranta, Finland
Phone number +358 40 8723522
Anna.Shcherbacheva@lut.fi

Abstract

Abstract

The aim of this work is to apply a nonlinear model from fluid mechanics to simulate electricity spot prices. The specified model is based on the viscous Burgers' equation, that provides the simplest mathematical model of turbulence in fluid dynamics. The model has been applied to Nord Pool spot market System and Pure Trading prices. Modification of the standart Kalman filter (VEnKF) is employed in order to adjust the model output. After carrying out forecasts it was observed that the model reacts upon the spikes afterward via rapid oscillation with large amplitudes.

Lappeenranta University of Technology
Department of Mathematics and Physics

Anna Shcherbacheva

Thesis for the Degree of Master of Science in Technology

year 2011

57 pages, 27 figures, 1 tables, 0 appendices

Examiners: prof. Ph.D. Tuomo Kauranne and prof. Ph.D. Heikki Haario

Keywords: Burgers' equation, Kalman filtering, statistical mechanics

Acknowledgements

Firstly, I would like to express my gratitude to the Department of Mathematics at Lappeenranta University of Technology for the scholarship during my studies, otherwise my studies would not be possible.

I am very grateful to my supervisor, Dr. Tuomo Kauranne, for this interesting topic for my master thesis and for his guidance during the whole process of carrying out a research.

And finally, I am very thankful to my family, to my husband Alexander , to my parents and grandmothers for their countenance throughout this year of studies.

Lappeenranta, July 15.

Anna Shcherbacheva

Contents

Abstract	i
Acknowledgements	ii
List of Tables	vi
List of Figures	vii
1 Introduction	1
2 Electricity spot markets	3
2.1 Nordic electricity spot market	3
2.2 Pure trading and regression model	4
2.2.1 Time Series Regression Model	4
2.2.2 Least Squares Estimation	5
2.2.3 Goodness of Fit	6
2.2.4 Pure Trading Series	6
3 Analogies between fluid mechanics and market dynamics	7
3.1 Diffusion equation	9
3.2 Capasso-Morale System of Stochastic Differential Equations .	10
4 One-dimensional fluid mechanics - the Burgers' equation	11

4.1	Topology of the domain	14
5	Numerical Methods for Burgers' equation	15
5.1	Finite Difference Schemes	16
5.1.1	Explicit finite difference scheme	16
5.1.2	Implicit Lax-Friedrichs scheme.	17
5.2	Wavelet-Galerkin Method	18
5.2.1	Burgers equation. Generalized solution.	18
5.2.2	Harmonic wavelets	19
5.2.3	Wavelet-Galerkin method	20
6	Research methods for Stochastic Differential Equations	21
6.1	The Euler–Maruyama Method	22
6.2	Milstein's Higher Order Method	24
6.3	Stochastic Runge-Kutta methods	25
7	Kalman Filters	26
7.1	Extended Kalman Filter	27
7.2	Variational Kalman Filter	29
7.3	Variational Ensemble Kalman Filter	30
8	A fluid mechanic model of Nordpool spot prices	31

9	Applying Kalman filtering to forecasting spot prices	35
10	Results	36
11	Discussion	52
	References	

List of Tables

1	Analogy between fluid mechanics and financial markets (partially adapted from [8]).	9
---	---	---

List of Figures

1	Plot of the regression model (green), real price (red) and pure price (blue) for a half year window (adapted from [17]). . . .	7
2	Characteristics and solution for Burgers' equation (small t) .	13
3	Shock formation in Burgers' equation	13
4	Real and imaginary parts of periodic harmonic wavelets: (a) $\psi_0^0(x)$ (b) $\psi_1^1(x)$ (c) $\psi_2^2(x)$ (d) $\psi_4^3(x)$	20
5	Real system price (blue) pure trading price (green) and their residual (red)	32
6	Ensemble mode (blue) and pure trading price (red) for ensemble size $N = 10$, $\alpha = 5/3, \theta = 0.08$	38
7	Ensemble mode (blue) and pure trading price (red) for ensemble size $N = 40$, $\alpha = 5/3, \theta = 0.08$	38
8	Ensemble mode (blue) and real system price (red) for ensemble size $N = 10$, $\alpha = 5/3, \theta = 0.08$	39
9	Relative error plot for real system price, $N = 10$, $\alpha = 5/3, \theta = 0.08$	39
10	Rmse error for real system price, $N = 10$, $\alpha = 5/3, \theta = 0.08$.	40
11	Rmse error for pure trading price for ensemble size $N = 10$, $\alpha = 5/3, \theta = 0.08$	40
12	Rmse error for pure trading price for ensemble size $N = 40$, $\alpha = 5/3, \theta = 0.08$	41

13	Forecast skills for Real system price for different ensemble size : N=10 (blue), N=15 (green), N=30 (yellow), N=40 (red) $\alpha = 5/3, \theta = 0.001$	41
14	Real price and 20 days ahead forecasts, $\alpha = 5/3$	42
15	Distributions of real price and 20 days ahead forecasts, $\alpha = 5/3$	44
16	Real price and 40 days ahead forecasts, $\alpha = 5/3$	44
17	Distributions of real price and 40 days ahead forecasts, $\alpha = 5/3$	45
18	Real price and 60 days ahead forecasts, $\alpha = 5/3$	45
19	Distributions of real price and 60 days ahead forecasts, $\alpha = 5/3$	46
20	Real price and 20 days ahead forecasts, $\theta = 0.08$	47
21	Distributions of real price and 20 days ahead forecasts, $\theta = 0.08$	48
22	Real price and 40 days ahead forecasts, $\theta = 0.08$	49
23	Distributions of real price and 40 days ahead forecasts, $\theta = 0.08$	50
24	Real price and 60 days ahead forecasts, $\theta = 0.08$	50
25	Distributions of real price and 60 days ahead forecasts, $\theta = 0.08$	51
26	Distributions of pure trading price and 20 days ahead forecasts, $\alpha = 5/3$	51
27	Distributions of pure trading price and 60 days ahead forecasts, $\alpha = 5/3$	52

1 Introduction

After introduction of competition and deregulation of the power markets the trading risk had abruptly increased due to extreme volatility of electricity prices. This is especially true for spot prices, where volatility can be as high as 50% on the daily scale, i.e. over ten times higher than for other energy products (natural gas and crude oil). High volatility enforces market participants to hedge against sharp price changes. In order to hedge efficiently, it is expedient to thoroughly study and, afterwards, accurately model electricity price dynamics. Designed models will make it possible to develop bidding strategies and negotiation skills to hedge producers and wholesale consumers and, simultaneously, to maximize their profits.

The main distinctive statistical feature of the electricity spot market price series is price spikes - the prices which within few hours can jump up to ten times higher from the mean level. Although market analysts are always able to explain the spike occurrence afterwards via some reasons, future spikes have seldom been produced by the same composition of factors, as electricity spot prices are simultaneously driven by both predetermined constituents, such as weather conditions, demand level and possibilities of energy production, as well as some unexpected random factors.

In recent years a considerable number of techniques have been developed to simulate electricity spot prices. Many of them adapt well-known stochastic models with addition of certain electricity price characteristics, like price spikes and mean reversion. Among the earliest models are jump-diffusion [15], and its extension - mean reverting jump-diffusion of Johnson and Barz [14]. A significant drawback of the latter model is the slow speed of mean reversion after a jump. However, this was subsequently overcome by adding downwards jumps, allowing time-varying parameters or incorporating nonlinearities in the price dynamics, such as regime switching and stochastic volatility [5], [12].

Since electricity spot prices exhibit excessive and non-constant volatility with evidence of heteroscedasticity both in unconditional and conditional variance, models that involve GARCH effects have been also employed. Karakatsani and Bunn [16] applied four approaches (including regression-GARCH and Time-Varying parameter regression with exogenous variables) to explain the stochastic dynamics of spot volatility and understand agent reactions of shocks. Further improvement was acquired when a regression model with the assumptions of an explicit jump component for prices and a leptocurtic distribution for innovations was introduced.

While such hybrid approaches incorporate some fundamental factors, the structural models are aimed at capturing more detailed patterns for electricity prices. For example, Vehviläinen and Pyykkönen [26] modelled hydrological inflow and snow-pack development that affect the main source of electricity in Scandinavia - hydro power generation. They provided a stochastic factor based approach, where the fundamentals affecting the spot price are modeled independently and then a market equilibrium model is used to combine them to determine spot prices. The model is rather designed for medium-term modeling and forecasting of electricity spot prices.

There are also universal approaches applied for price predictions, like artificial intelligence approaches based on neural networks and fuzzy logic [25], [28]. However, the main disadvantages of artificial intelligence based techniques are a lack of intuitiveness and the fact that no simple physical interpretation is possible for their components.

In this study we investigate the applicability of fluid mechanics and Kalman filtering to forecasting electricity spot prices. This work is divided into several sections. The first section provides an introduction to electricity spot markets, in the second one some analogues between fluid mechanics and market dynamics are described. The fourth and the fifth sections are devoted to the Burgers' equation and the numerical methods for its resolution, correspondingly. Theoretical background on methods for Stochastic Differential

Equations have been represented in section six. Section seven is devoted to Kalman filtering in several modifications (EKF, VKF and VEnKF).

A detailed description of a fluid mechanic model of Nord Pool spot prices has been given in section eight, while the subject of the ninth section is Kalman filtering applied to forecasting the Nord Pool spot prices. Two final sections present practical results and discuss conclusions that can be arrived at from them.

2 Electricity spot markets

2.1 Nordic electricity spot market

The Nordic electricity exchange Nord Pool is the world's only multinational exchange for trading electric power, operating in Norway, Denmark, Sweden, Finland and Estonia. The company represents over 70% of the total electricity consumption in the Nordic countries. Since 2005 Nord Pool provides also emission allowances and emission credits.

The Nord Pool Spot market membership includes energy producers, energy intensive industries, large consumers, distributors, funds, investment companies, banks, brokers, utility companies and financial institutions [34].

Nord Pool regulates the market in an optimal way, matching electricity generation supply with demand on an hourly basis. The trading cycle is organized as a closed auction and it is conducted once a day. By noon of a given day the market administrator receives the electronically sent bids and offers from producers and consumers for each hourly period of the next day and calculates hourly spot prices as an equilibrium point of supply and demand for each particular hour.

There are three possible ways of bidding at the day-ahead market Elspot. The first kind of bidding - hourly bidding - involves a pair of price and volume for each hour, while for the second kind the price and the volume

remain identical for a number of successive hours. Lastly, flexible bidding implies fixed price and volume with flexible hour of the sale determined by the highest day price that is above the price specified by the bid.

In case of grid congestions, the Nordic Exchange area is divided into pre-determined price areas that are separated by congested transmission lines. Inside of the price areas, congestions are not expected to happen. Within the Nordic electricity market, Norway is subdivided into five price areas, Denmark composes two price areas, while Sweden, Finland and Estonia constitute one price area each[13]. The participants' bids in the bidding areas on each side of the congestion is aggregated into supply and demand curves in the same way as in the system price calculation [33].

2.2 Pure trading and regression model

2.2.1 Time Series Regression Model

Consider the *linear time series regression model*

$$Y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t = \mathbf{x}'_t \boldsymbol{\beta} + \varepsilon_t, t = 1, \dots, T \quad (2.2.1)$$

where $\mathbf{x}_t = (1, x_{t1}, x_{t2}, \dots, x_{tk})'$ is a $(k + 1) \times 1$ vector of the **background variables**, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)'$ is a $(k + 1) \times 1$ vector of unknown parameters to be estimated, and ε_t is a random error term. The part of the right-hand side involving the regressors, $\beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t$, is called the **regression** or the **regression function**, and the coefficients (β 's) are called the **regression coefficients** [10].

In matrix form the model (2.2.1) is expressed as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.2.2)$$

where $\mathbf{y}, \boldsymbol{\varepsilon}$ are $T \times 1$ vectors and \mathbf{X} is $T \times (k + 1)$ matrix. The standard assumptions of the time series regression model are

- the linear model (2.2.1) is correctly specified.
- $\{y_t, \mathbf{x}_t\}$ is jointly stationary and ergodic.
- the regressors \mathbf{x}_t are *predetermined*: $E[x_{is}\varepsilon_t] = 0$ for all $s \leq t$ and $i = 1, \dots, k$.
- $E[\mathbf{x}_t \mathbf{x}_t'] = \Sigma_{\mathbf{X}\mathbf{X}}$ is of full rank $k + 1$.
- $\{\mathbf{x}_t, \varepsilon_t\}$ is an uncorrelated process with finite $(k+1) \times (k+1)$ covariance matrix $E[\varepsilon_t^2 \mathbf{x}_t \mathbf{x}_t'] = \mathbf{S} = \sigma^2 \Sigma_{\mathbf{X}\mathbf{X}}$.

The second assumption rules out trending regressors, the third rules out endogenous regressors but allows lagged dependent variables.

To understand the fourth assumption, recall from matrix algebra that the rank of a matrix equals the number of linearly independent columns of the matrix. The assumption says that none of the $k + 1$ columns of the data matrix \mathbf{X} can be expressed as a linear combination of the other columns of \mathbf{X} . That is, \mathbf{X} is of **full column rank**. Since the $k + 1$ columns cannot be linearly independent if their dimension is less than $k + 1$, the assumption implies that $T \geq k + 1$, i.e., there must be at least as many observations as there are regressors [10].

The fifth assumption implies that error term is a serially uncorrelated process with constant unconditional variance σ^2 . In the time series regression model, the regressors \mathbf{x}_t are random and the error term ε_t is not assumed to be normally distributed.

2.2.2 Least Squares Estimation

Ordinary least squares (OLS) estimation is based on minimizing the sum of squared residuals

$$SSR(\boldsymbol{\beta}) = \sum_{t=1}^T (y_t - \mathbf{x}_t' \boldsymbol{\beta})^2 = \sum_{t=1}^T \varepsilon_t^2 \quad (2.2.3)$$

and produces the fitted model

$$y_t = \mathbf{x}'_t \hat{\boldsymbol{\beta}} + \hat{\varepsilon}_t, t = 1, \dots, T \quad (2.2.4)$$

where $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $\hat{\varepsilon}_t = y_t - \hat{y}_t = y_t - \mathbf{x}'_t \hat{\boldsymbol{\beta}}$. The error variance is estimated as $\hat{\sigma}^2 = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}/(T - k - 1)$. Under the assumptions described above, the OLS estimates $\hat{\boldsymbol{\beta}}$ are consistent and asymptotically normally distributed [30].

2.2.3 Goodness of Fit

Goodness of fit is summarized by the R of the regression

$$R = 1 - \frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{(\mathbf{y} - \bar{y}\mathbf{1})'(\mathbf{y} - \bar{y}\mathbf{1})} \quad (2.2.5)$$

where \bar{y} is the sample mean of y_t and $\mathbf{1}$ is a $T \times 1$ vector of 1's. R^2 measures the percentage of the variability of y_t that is explained by the regressors, x_t . The usual R^2 has the undesirable feature of never decreasing as more variables are added to the regression, even if the extra variables are irrelevant.

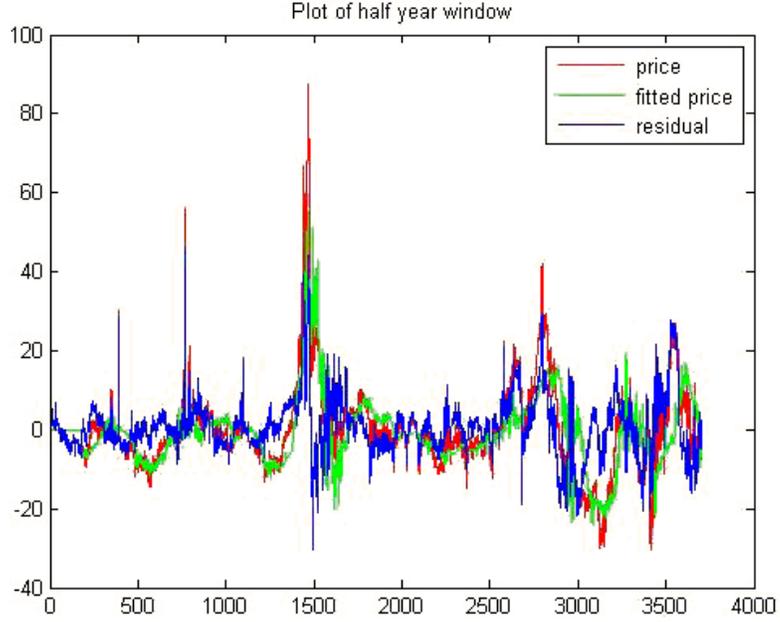
2.2.4 Pure Trading Series

By pure-trading we mean the price series (behaviour) with eliminated trend and seasonality impact. The pure price represents a general systematic non-linear component that changes over time and does not repeat or at least does not repeat within the time range captured by the data.

In a recent study [17] a regression model was applied to pure-trading series, employing hydrological storage level and temperature as the background variables. Before estimating the desired model, the explanatory variables were also detrended and deseasonalized to comply with the dependent variable. Finally, the least-squares-optimal regression model was built in a moving regression fashion with half-a-year history window, see Figure 1.

In order to obtain a better fitting, there should be more background variables taken into account, e.g. prices of fossil fuels, rainfall and non-base demand [4].

Figure 1: Plot of the regression model (green), real price (red) and pure price (blue) for a half year window (adapted from [17]).



3 Analogies between fluid mechanics and market dynamics

The Navier-Stokes equations, describing the motion of an incompressible fluid are given by:

$$\begin{cases} \partial_t \vec{u} + \vec{u} \cdot \nabla \vec{u} = -\nabla p + \nu \Delta \vec{u} \\ \nabla \cdot \vec{u} = 0 \end{cases} \quad (3.1)$$

$u(\vec{r}, t)$ denotes the velocity of the fluid at position \vec{r} and time t and p is the pressure, the left hand side describes the total derivate of the velocity.

The control parameter, which governs the behaviour of a liquid, is the dimensionless *Reynolds number* R :

$$R = \frac{Lu}{\nu} \quad (3.2)$$

where L is a typical length scale of the observed system, ν is the kinematic viscosity. R is a measure for the complexity of the flowing fluid, $\vec{u} = u\vec{f}(\vec{r}/L, t)$. For small R the motion of the fluid is laminar. When R becomes bigger the flow turns into a turbulent flow.

A flow of energy from large to small scales is one of the main characteristics of a fully developed turbulence. It results in dissipation of large amounts of energy in a viscous fluid. Kolmogorov and Obhukov stated that when energy is injected in big eddies that they break up into smaller eddies until the energy is dissipated on the smallest scale. This cascade of kinetic energy results in a scaling of the moments $\langle(\Delta u)^n\rangle$ of Δu as $(\Delta r)^{\xi_n}$, where the angle brackets denote the mean value of the inclosed quantity and Δu is the difference of the velocity component in the direction of the spatial separation of length Δr , see [8].

On a qualitative level financial markets and turbulence have a tempting similarity. The analogy to energy in turbulence is said to be information in financial markets [27]. Müller et al. [21] found that there are two different types of traders. The long-term traders evaluate the market on a low frequency and have a long memory. The short-term traders watch the market continuously, re-evaluate the situation and execute transactions more often but they have a shorter memory. That is why they cause different types of volatility. On a short time scale the volatility is dominated by the actions of the short-time traders, while the long-term traders cause the volatility on a long time grid. Furthermore they found, that the information flow between these two different classes of traders is asymmetric. The short-term traders react on peaks of the long-term volatility with increased trading activity, but the long-term traders often ignore the fluctuations on short time scales. This

net information flow from long to short time scales is said to be the *information cascade* in financial markets. The claimed analogy between different quantities in fluid mechanics and financial markets is summarized in Table 1.

Table 1: Analogy between fluid mechanics and financial markets (partially adapted from [8]).

Hydrodynamic turbulence	Financial markets
Energy	Information
Spatial distance	Time delay
Laminar periods interrupted by turbulent bursts (intermittency)	Clusters of low and high volatility
Energy cascade in space hierarchy	Information cascade in time hierarchy
Advection of the particles	Traders' movement towards the higher price
$\langle (\Delta u)^n \rangle \propto (\Delta r)^{\xi_n}$	$\langle (\Delta x)^n \rangle \propto (\Delta t)^{\xi_n}$

3.1 Diffusion equation

Assume that prices are quantized into classes $\dots, S_{n-2}, S_{n-1}, S_n, S_{n+1}, S_{n+2}, \dots$ and that at some time t in the future, these prices are realized with probabilities $\dots, p_{n-2}, p_{n-1}, p_n, p_{n+1}, p_{n+2}, \dots$. Then, one may ask for the evolution of these probabilities with time. If we assume that a price change $S_n \rightarrow S_{n\pm 1}$ must take place during Δt , the probability p'_n of having S_n at time step Δt after t will be $p'_n = (p_{n-1} + p_{n+1})/2$ because the price S_n can either be reached by a downward move from S_{n+1} , occurring with a probability $p_{n+1}/2$, or by an upward move from S_{n-1} with a probability $p_{n-1}/2$. The change in probability of a price S_n during the time step Δt is then

$$\Delta p_n = p'_n - p_n = \frac{p_{n-1} + 2p_n - p_{n+1}}{2} \rightarrow \frac{\partial^2 p(S, t)}{\partial S^2} (\Delta S)^2 \quad (3.1.1)$$

if the limit of continuous prices and time is taken. On the other hand,

$$\Delta p_n \rightarrow \frac{\partial p(S, t)}{\partial t} \Delta t \quad (3.1.2)$$

in the same limit, and therefore

$$D \frac{\partial^2 p}{\partial S^2} - \frac{\partial p}{\partial t} = 0 \quad (3.1.3)$$

Hence, $p(S, t)$ satisfies a diffusion equation, describing the motion of suspended particles in a solvent - the so-called *Brownian motion*. In case of special initial conditions $p(S, 0) = \delta[S - S(0)]$, i.e., knowledge of the price at time $t = 0$ the Gaussian distribution is obtained [27].

This reveals, that the diffusion equation corresponds to the *Efficient Market Hypothesis (EMH)*. The EMH states that the current price of an asset fully reflects all available information relevant to it and that new information is immediately incorporated into the price, while the future flow of news (that will determine future stock prices) is random and unknowable in the present. It also assumes that there is a rational and unique way to use available information, that all agents possess this knowledge and that any chain reaction produced by a ‘shock’ happens instantaneously [32].

3.2 Capasso-Morale System of Stochastic Differential Equations

Opponents of the EMH sometimes cite examples of market movements that seem inexplicable in terms of conventional theories of stock price determination, for example the stock market crash of October 1987 where most stock exchanges crashed at the same time. The correct explanation seems to lie either in the mechanics of the exchanges (e.g. no safety nets to discontinue trading initiated by program sellers) or the peculiarities of human nature [32]. An alternative theory, prominent as the Behavioral Finance, attempts to build models accounting for human behaviour and psychology. In Financial market the latter manifests itself mostly in fear and greed, influenced then by our common trading biases: herding, overconfidence and short-term thinking.

A fruitful approach, based on a Capasso-Morale type system of stochastic differential equations, is based on the modelling of the behavior of each individual trader in a group of traders in the spot market. It was used so far for modelling animal population dynamics. Therefore, an evolution of each individual price is driven by the system of SDE of the type:

$$dX_N^k(t) = [f_N^k + h_N^k(X_N^1(t), \dots, X_N^N(t), t)]dt + \sigma(X_N^1(t), \dots, X_N^N(t), t)dW^k(t), \quad k = 1, \dots, N \quad (3.2.1)$$

where the function $f_N^k : \mathbb{R}_+ \rightarrow \mathbb{R}$ describes the individual dynamics which may depend only on time or on the state of the particle itself, $h_N^k : \mathbb{R}^N \times \mathbb{R}_+$ stands for the interaction of the k-th trader with other traders in the group. Here an additive noise is modelled via a family of independent Wiener processes.

It is assumed that traders do observe one another and thus tend to follow a general price path, like in model, described by the diffusion equation. However, there is a limit for overcrowding (in microscale) which in power trading could be interpreted as physical impossibility of two market participants to buy the same dose of electricity [20]. The force that describes the individual dynamics corresponds to the momentum term in Navier-Stokes equations: e.i., it describes the fact, that the individual prices deviate from the global tendency due to the traders pursuit of an extra profit.

4 One-dimensional fluid mechanics - the Burgers' equation

Burgers' equation is a fundamental partial differential equation from fluid mechanics. It occurs in various areas of applied mathematics, such as modelling gas dynamics and traffic flow. It is named for Johannes Martinus Burgers (1895-1981). Since this single equation has a convective term, a

diffusive term and a time-dependent term, it can serve as a simplest model that includes the nonlinear and viscous effects of fluid dynamics.

It has been discussed by Burgers as a mathematical model of turbulence and by Cole as the approximate theory for weak non-stationary shock waves in a real fluid. This equation provides an approximation of the Navier-Stokes equations in the case where the net force acts mostly in one direction.

For a given velocity u and viscosity coefficient ν , the general form of viscous Burgers' equation is:

$$u_t + \alpha uu_x + \nu u_{xx} = 0 \quad (4.1)$$

Burgers' equation is parabolic when the viscous term is included. If the viscous term is neglected, the remaining equation is hyperbolic:

$$u_t + uu_x = 0 \quad (4.2)$$

If the viscous term is dropped from the Burgers' equation the nonlinearity allows discontinuous solutions to develop. Consider the inviscid equation 4.2 with smooth initial data. For small time t , a solution can be constructed by following characteristics:

$$x'(t) = u(x(t), t) \quad (4.3)$$

and along each characteristic u is constant, since

$$\begin{aligned} \frac{d}{dt}u(x(t), t) &= \frac{\partial}{\partial t}u(x(t), t) + \frac{\partial}{\partial x}u(x(t), t)x'(t) \\ &= u_t + uu_x \\ &= 0 \end{aligned} \quad (4.4)$$

Moreover, since u is constant on each characteristic, the slope $x'(t)$ is constant by 4.3 and so the characteristics are straight lines, determined by the initial data (see Figure 2).

If the initial data is smooth, then this can be used to determine the solution $u(x, t)$ for small enough t that characteristics do not cross. For each (x, t) we can solve the equation

$$x = \xi + u(\xi, 0)t \quad (4.5)$$

for ξ and then

$$u(x, t) = u(\xi, 0) \quad (4.6)$$

Figure 2: Characteristics and solution for Burgers' equation (small t)

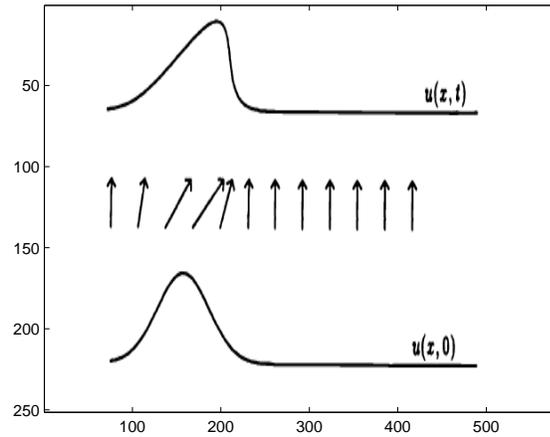
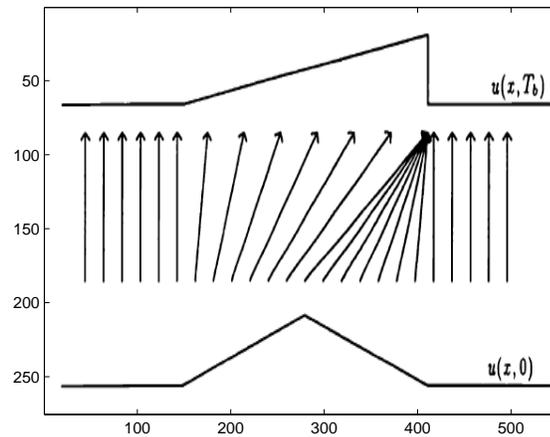


Figure 3: Shock formation in Burgers' equation



For larger t the equation 4.5 may not have a unique solution. This happens when the characteristics cross. At the time T_b where the characteristics first cross, the function $u(x, t)$ has an infinite slope - the wave "breaks" and a

shock forms. Beyond this point there is no classical solution of the PDE, and the weak solution becomes discontinuous. Figure 3 shows an extreme example where the initial data is piecewise linear and many characteristics come together at once, [18].

Consider the viscous Burgers' equation with an external forcing:

$$u_t + \alpha uu_x + \nu u_{xx} = f(x, t) \quad (4.7)$$

Describing 4.7 in terms of market dynamics, we come up with the following analogies:

- u specifies price,
- $f(x, t)$ stands for fundamentals (i.e., external forces of periodic character affecting system dynamic),
- νu_{xx} - the diffusion term describing the fact that the spot market tends to reach the equilibrium price,
- αuu_x - the momentum term expressing traders' movement towards higher price,
- u_x - the spread of bids for a given hour or day.

4.1 Topology of the domain

The Nord Pool spot market can be considered as a group of traders, observing one another, so that if one trader has changed the price significantly, the bidding of neighbouring participants will strongly correlate with this event. Reciprocally, their neighbours will take into account observed changes while making their decisions, and so forth. Finally, after the splash propagates throughout the whole group, attenuating subsequently, the market feedback will influence the initial trader that had caused the chain reaction. This

assumptions can be introduced via choosing a periodic topological domain. Since there are neither inflow nor outflow constraints, advected waves are able to propagate fluently through the boundaries.

However, it is more reasonable to consider an infinite domain, since in reality the Elspot participants observe one another within the small microgroups, so that in most cases the traders from inside are not strongly influenced by the external changes. This particular choice of the topological domain imposes the advected perturbations to have limited range of outspread, simultaneously excluding the possibility of market feedback. The latter effect can be taken into account by considering the semi-infinite case and specifying an appropriate boundary condition, so that the energy won't be neither absorbed nor superinduced at the boundary.

5 Numerical Methods for Burgers' equation

Further we consider the following initial-boundary value problem for Burgers' equation:

$$u_t + \alpha uu_x + \vartheta u_{xx} = f(x, t), \quad 0 \leq x \leq 1 \quad (5.1)$$

$$u(x, 0) = \phi(x) \quad (5.2)$$

$$u(0, t) = u(1, t) \quad (5.3)$$

We describe explicit and implicit finite difference schemes for above-stated IBV problem and then try to proceed in a different way using Wavelet-Galerkin method (see [35]).

5.1 Finite Difference Schemes

We discretize the $x - t$ plane by choosing a mesh width $h \equiv \Delta x$ and a time step $k \equiv \Delta t$ and define the discrete mesh points (x_i, t_n) by.

$$\begin{aligned} x_i &= ih, \quad i = 0, 1, \dots, M \\ t_n &= nk, \quad n = 0, 1, \dots, N \end{aligned} \quad (5.1.1)$$

Where,

$$M + 1 = 1/h, \quad \text{and } N = T/k \quad (5.1.2)$$

5.1.1 Explicit finite difference scheme

To obtain an explicit finite difference scheme, we discretize $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x}$ and $\frac{\partial^2 u}{\partial x^2}$ at any mesh point (x_i, t_n) as follows:

$$\frac{\partial u}{\partial t} \approx \frac{u_i^{n+1} - u_i^n}{k} \quad (\text{First order Forward difference formula}) \quad (5.1.3)$$

$$\frac{\partial u}{\partial x} \approx \frac{u_{i+1}^n - u_{i-1}^n}{2h} \quad (\text{First order Centered difference formula}) \quad (5.1.4)$$

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \quad (\text{Second order Centered difference formula}) \quad (5.1.5)$$

Taking into account for boundary conditions and inserting 5.1.3, 5.1.4, 5.1.5 in 5.1, 5.3 the discrete version of the IBV problem 5.1-5.3 formulates the second order finite difference scheme of the form:

$$\frac{u_0^{n+1} - u_0^n}{k} + \alpha \frac{u_0^n (u_1^n - u_M^n)}{2h} + \vartheta \frac{u_1^n - 2u_0^n + u_M^n}{h^2} = f_0^n, \quad (5.1.6)$$

$$\begin{aligned} \frac{u_i^{n+1} - u_i^n}{k} + \alpha \frac{u_i^n (u_{i+1}^n - u_{i-1}^n)}{2h} + \vartheta \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} \\ = f_i^n, \quad i = 1, \dots, M - 1, \end{aligned} \quad (5.1.7)$$

$$\frac{u_M^{n+1} - u_M^n}{k} + \alpha \frac{u_M^n (u_0^n - u_{M-1}^n)}{2h} + \vartheta \frac{u_0^n - 2u_M^n + u_{M-1}^n}{h^2} = f_M^n \quad (5.1.8)$$

$$u_i^0 = \phi_i, \quad i = 0, \dots, M \quad (5.1.9)$$

which is the explicit finite difference scheme for IBVP:

$$u_0^{n+1} = u_0^n - \frac{\alpha k u_0^n}{2h} (u_1^n - u_M^n) - \frac{\vartheta k}{h^2} (u_1^n - 2u_0^n + u_M^n) + k f_0^n, \quad (5.1.10)$$

$$u_i^{n+1} = u_i^n - \frac{\alpha k u_i^n}{2h} (u_{i+1}^n - u_{i-1}^n) - \frac{\vartheta k}{h^2} (u_{i+1}^n - 2u_i^n + u_{i-1}^n) + k f_i^n, \quad i = 1, \dots, M-1, \quad (5.1.11)$$

$$u_M^{n+1} = u_M^n - \frac{\alpha k u_M^n}{2h} (u_0^n - u_{M-1}^n) - \frac{\vartheta k}{h^2} (u_0^n - 2u_M^n + u_{M-1}^n) + k f_M^n, \quad (5.1.12)$$

$$u_i^0 = \phi_i, \quad i = 0, \dots, M \quad (5.1.13)$$

However, it is more common to discretize Burgers' equation in the so-called **conservative form**, since according to the **Lax-Wendroff** theorem, if a solution of a conservative scheme converges ($\Delta x \rightarrow 0$), it converges towards a weak solution of the conservation law. In **conservative form** Burgers' equation takes the form:

$$u_t + \alpha F_x + \vartheta u_{xx} = f(x, t), \quad F = \frac{u^2}{2} \quad (5.1.14)$$

The explicit Lax-Friedrichs scheme for conservative formulation 5.1.14 can be written in the following way:

$$u_i^{n+1} = \frac{u_{i+1}^n + u_{i-1}^n}{2} - \frac{\alpha k}{4h} ((u_{i+1}^n)^2 - (u_{i-1}^n)^2) - \frac{\vartheta k}{h^2} (u_{i+1}^n - 2u_i^n + u_{i-1}^n) + k f_i^n \quad (5.1.15)$$

It is interesting to note that the explicit treatment of the viscous term has been shown not to converge, while the viscous term treated implicitly makes Lax-Friedrichs scheme stable, [1].

5.1.2 Implicit Lax-Friedrichs scheme.

Consider the following discretization of the Burgers' equation:

$$u_i^{n+1} = \frac{u_{i+1}^n + u_{i-1}^n}{2} - \frac{\alpha k}{4h} ((u_{i+1}^n)^2 - (u_{i-1}^n)^2) - \frac{\vartheta k}{h^2} (u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}) + k f_i^n, \quad i = 1, \dots, M-1. \quad (5.1.16)$$

Theorem 1 Denote an exact solution of original IBVP as z , let us define for any vector $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$,

$$\|x\|_\infty = \max_{1 \leq j \leq n} \{|x_j|\} \quad (5.1.17)$$

and for the bounded function z defined on the domain $I \times [0, T]$,

$$\|z\|_\infty = \sup_{I \times [0, T]} \{|z|\} \quad (5.1.18)$$

If the condition 5.1.19

$$\frac{kC_0}{h} \leq 1, \quad (5.1.19)$$

where $C_0 = \max\{\|\phi\|_\infty, \|z\|_\infty\}$, holds for k, h with $kM \leq T$ and $\frac{k}{h}$ is a positive constant, the implicit Lax-Friedrichs scheme 5.1.16 is convergent, [1].

5.2 Wavelet-Galerkin Method

5.2.1 Burgers equation. Generalized solution.

In the present method the boundary conditions are automatically taken into account due to proper choice of the basic functions. Here instead of the classical we consider a generalized solution such that $u(x, \cdot) \in \hat{W}_2^1(0, 1)$, $u(\cdot, t) \in C^0(\mathbb{R})$, $u_t(\cdot, t) \in L_2(0, 1)$, where $\hat{W}_2^1(0, 1)$ is a Sobolev space of integrable functions

$$\begin{aligned} \hat{W}_2^1(0, 1) &= \{f \in L_2(0, 1) \mid \exists D^1 f \in L_2(0, 1), \\ \|f\|_{W_2^1(0, 1)} &= (\|f\|_{L_2(0, 1)} + \|D^1 f\|_{L_2(0, 1)})^{\frac{1}{2}}, f(0) = f(1)\} \end{aligned} \quad (5.2.1)$$

Employing the fact that $\overline{C^\infty}(0, 1)$ is densely embedded into $W_2^1(0, 1)$, we define generalized solution as a function u , such that for every smooth test

function $\varphi(x) \in C^\infty(0, 1)$, $\varphi(0) = \varphi(1)$ the following equality holds:

$$\begin{aligned} & \int_0^1 u_t(x, t) \cdot \varphi(x) dx + \int_0^1 \alpha \left(\frac{u^2(x, t)}{2} \right)_x \cdot \varphi(x) dx \\ & + \int_0^1 \vartheta u_{xx}(x, t) \cdot \varphi(x) dx = \int_0^1 f(x, t) \cdot \varphi(x) dx \end{aligned} \quad (5.2.2)$$

Applying integration by parts to move derivatives from dependent variable on the test function and taking into account for boundary conditions we ultimately define the equality fulfilled for required solution:

$$\begin{aligned} & \int_0^1 u_t(x, t) \cdot \varphi(x) dx - \frac{\alpha}{2} \int_0^1 u^2(x, t) \cdot \varphi_x(x) dx + \\ & \vartheta \int_0^1 u(x, t) \cdot \varphi_{xx}(x) dx = \int_0^1 f(x, t) \cdot \varphi(x) dx, \\ & \forall \varphi(x) \in C^\infty(0, 1), \varphi(0) = \varphi(1) \end{aligned} \quad (5.2.3)$$

Otherwise, expressing in terms of scalar product we constitute:

$$\begin{aligned} \langle u_t(\cdot, t), \varphi \rangle_{L_2(0,1)} - \frac{\alpha}{2} \langle u^2(\cdot, t), \varphi_x \rangle_{L_2(0,1)} + \vartheta \langle u(\cdot, t), \varphi_{xx} \rangle_{L_2(0,1)} \\ = \langle f(\cdot, t), \varphi \rangle_{L_2(0,1)}, \forall \varphi \in C^\infty \end{aligned} \quad (5.2.4)$$

5.2.2 Harmonic wavelets

Let us consider a complex Littlewood-Paley basis defined from the mother wavelet:

$$\psi(x) = \frac{e^{4i\pi x} - e^{2i\pi x}}{2i\pi x} \quad (5.2.5)$$

Applying translation and extension consequently, the orthonormal basic functions are constructed:

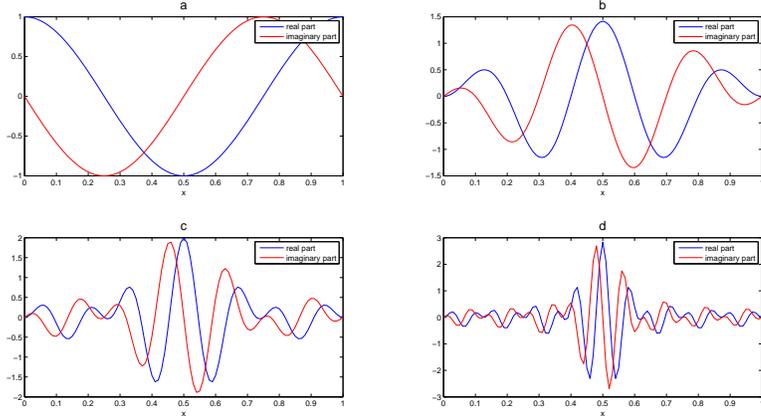
$$\psi_j^k(\omega) = \begin{cases} \frac{e^{i\omega k/2^j}}{2\pi \cdot 2^{j/2}} & 2\pi \cdot 2^j < 2\pi\omega \cdot 2^{j+1} \\ 0 & \text{otherwise} \end{cases} \quad (5.2.6)$$

and finally, transformed to periodical analogues at the unit interval:

$$\psi_j^{k \text{ per}}(x) = \frac{1}{2^{j/2}} \sum_{m_j=2^j}^{2^{j+1}-1} e^{-2i\pi(x-k/2^j)m_j} \quad (5.2.7)$$

Figure 4 contains both real and imaginary parts of periodic harmonic wavelets $\psi_0^0(x)$, $\psi_1^1(x)$, $\psi_2^2(x)$ and $\psi_4^3(x)$.

Figure 4: Real and imaginary parts of periodic harmonic wavelets: (a) $\psi_0^0(x)$ (b) $\psi_1^1(x)$ (c) $\psi_2^2(x)$ (d) $\psi_4^3(x)$



5.2.3 Wavelet-Galerkin method

In this subsection harmonic wavelets are employed to obtain a generalized solution satisfying the equation:

$$\begin{aligned} \langle u_t(\cdot, t), \varphi \rangle_{L_2(0,1)} - \frac{\alpha}{2} \langle u^2(\cdot, t), \varphi_x \rangle_{L_2(0,1)} + \vartheta \langle u(\cdot, t), \phi_{xx} \rangle_{L_2(0,1)} \\ = \langle f(\cdot, t), \phi \rangle_{L_2(0,1)}, \forall \phi \in C^\infty \end{aligned} \quad (5.2.8)$$

Here the brackets $\langle \cdot, \cdot \rangle$ express the scalar product:

$$\langle f, g \rangle_{L_2(0,1)} = \int_0^1 f(x) \bar{g}(x) dx, \quad (5.2.9)$$

where an overline denotes a complex conjugate. Using a standard approach for the Galerkin method, we introduce $u(x, t)$ via combination of the time-dependent amplitude and the basis, depending on the spatial variable:

$$u(x, t) = \sum_{j=0}^r \sum_{k=0}^{2^j-1} a_k^j(t) \psi_k^j(x) \quad (5.2.10)$$

Substituting 5.2.10 into 5.2.8, we obtain equation with respect to the time-dependent coefficients $a_s^r(t)$:

$$\begin{aligned} \frac{d}{dt} a_s^r(t) - \frac{\alpha}{2} \sum_{j=0}^r \sum_{p=0}^r \sum_{k=0}^{2^j-1} \sum_{q=0}^{2^j-1} N_{qks}^{jpr} a_k^j(t) a_q^p(t) \\ + \vartheta \sum_{j=0}^r \sum_{k=0}^{2^j-1} L_{ks}^{jr} a_k^j(t) = f_s^r(t) \end{aligned} \quad (5.2.11)$$

where:

$$N_{qks}^{jpr} = \langle \psi_k^j \frac{d}{dx} \psi_q^p, \psi_s^r \rangle \quad (5.2.12)$$

$$L_{ks}^{jr} = \langle \frac{d^2}{dx^2} \psi_k^j, \psi_s^r \rangle \quad (5.2.13)$$

$$f_s^r(t) = \langle f(\cdot, t), \psi_s^r \rangle \quad (5.2.14)$$

Thereafter, by swapping summation with integration and calculating the internal integrals we finally express coefficients 5.2.12, 5.2.13 explicitly:

$$N_{qks}^{jpr} = -\frac{2i\pi}{2^{\frac{j+p+r}{2}}} \sum_{m_j=2^j}^{2^{j+1}-1} \sum_{m_p=2^p}^{2^{p+1}-1} \sum_{m_r=2^r}^{2^{r+1}-1} m_p e^{2i\pi \left(\frac{m_j k}{2^j} - \frac{m_p q}{2^p} - \frac{m_r s}{2^r} \right)} \delta_{m_j+m_p, m_r} \quad (5.2.15)$$

$$L_{ks}^{jr} = -\frac{4\pi^2}{2^{(j+r)/2}} \sum_{m_j=2^j}^{2^{j+1}-1} \sum_{m_r=2^r}^{2^{r+1}-1} m_j^2 e^{2i\pi \left(\frac{m_j k}{2^j} - \frac{m_r s}{2^r} \right)} \delta_{m_j, m_r} \quad (5.2.16)$$

6 Research methods for Stochastic Differential Equations

Brownian Motion. A scalar *standard Brownian motion*, or *standard Wiener process*, over $[0, T]$ is a random variable $W(t)$ that depends continuously on

$t \in [0, T]$ and satisfies the following three conditions:

1. $W(0) = 0$ (with probability 1).
2. For $0 \leq s < t \leq T$ the random variable given by the increment $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$, equivalently, $W(t) - W(s) \sim \sqrt{t - s}N(0, 1)$.
3. For $0 \leq s < t < u < v \leq T$ the increments $W(t) - W(s)$ and $W(v) - W(u)$ are independent.

Stochastic integrals. Given a suitable function h , consider a sum of the form

$$\sum_{j=0}^{N-1} h(t_j)(W(t_{j+1}) - W(t_j)) \quad (6.1)$$

which by analogy with the approximation of the Riemann integral, may be regarded as an approximation to a stochastic integral $\int_0^T h(t)dW(t)$. Here, we are integrating h with respect to Brownian motion.

An alternative to 6.1 is given by

$$\sum_{j=0}^{N-1} h\left(\frac{t_j + t_{j+1}}{2}\right)(W(t_{j+1}) - W(t_j)) \quad (6.2)$$

The “left-hand” sum 6.1 gives in limit (when $\delta t \rightarrow 0$) what is known as the *Itô* integral, whereas the “midpoint” sum 6.2 produces the *Stratonovich* integral.

6.1 The Euler–Maruyama Method

Consider a scalar autonomous SDE

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad X(0) = X_0, \quad 0 \leq t \leq T \quad (6.1.1)$$

Here, f and g are scalar functions and the initial condition X_0 is a random variable. The solution $X(t)$ is a random variable for each t .

We assume that f and g are defined and measurable in $[0, T] \times \mathbb{R}$ and satisfy both Lipschitz and linear growth bound conditions. These requirements ensure existence and uniqueness of solution of the SDE 6.1.1.

If $g \equiv 0$ and X_0 is constant, then the problem becomes deterministic, and 6.1.1 reduces to the ordinary differential equation $dX(t)/dt = f(X(t))$, with $X(0) = X_0$.

SDE 6.1.1 can be rewritten in integral form as

$$X(t) = X(0) + \int_0^t f(X(s))ds + \int_0^t g(X(s))dW(s), \quad 0 \leq t \leq T \quad (6.1.2)$$

The second integral on the right-hand side of 6.1.2 is to be taken with respect to Brownian motion. To apply a numerical method to 6.1.1 over $[0, T]$, we first discretize the interval. Let $\Delta t = T/L$, where L is some positive integer, and $\tau_j = j\Delta t$. Our numerical approximation to $X(\tau_j)$ will be denoted as X_j . The Euler–Maruyama (EM) method takes the form

$$\begin{aligned} X_j = & X_{j-1} + f(X_{j-1})\Delta t \\ & + g(X_{j-1})(W(\tau_j) - W(\tau_{j-1})), \quad j = 1, 2, \dots, L \end{aligned} \quad (6.1.3)$$

Notice from the integral form 6.1.2, that

$$X(\tau_j) = X(\tau_{j-1}) + \int_{\tau_{j-1}}^{\tau_j} f(X(s))ds + \int_{\tau_{j-1}}^{\tau_j} g(X(s))dW(s) \quad (6.1.4)$$

Each of the three terms on the right-hand side of 6.1.3 approximates the corresponding term on the right-hand side of 6.1.4. It is also should be mentioned that in the deterministic case ($g \equiv 0$ and X_0 is constant), 6.1.3 reduces to Euler’s method.

Strong and weak convergence. A method is said to have *strong order of convergence* equal to γ if there exists a constant C such that

$$\mathbb{E}|X_n - X(\tau)| \leq C\Delta t^\gamma \quad (6.1.5)$$

for any fixed $\tau = n\Delta t \in [0, T]$ and Δt sufficiently small. Here \mathbb{E} denotes the expected value. If f and g satisfy appropriate conditions, it can be shown

that EM has strong order of convergence $\gamma = \frac{1}{2}$. Note that in deterministic case ($g \equiv 0$ and X_0 is constant) the expected value can be deleted from the left-hand side of 6.1.5 and the inequality is true with constant $\gamma = 1$.

The strong order of convergence measures the rate at which the “mean of the error” decays as $\Delta t \rightarrow 0$. A less demanding alternative is to measure the rate of decay of the “error of the means.” This leads to the concept of *weak convergence*. A method is said to have weak order of convergence equal to γ if there exists a constant C such that for all functions p in some class

$$|\mathbb{E}p(X_n) - \mathbb{E}p(X(\tau))| \leq C\Delta t^\gamma \quad (6.1.6)$$

at any fixed $\tau = n\Delta t \in [0, T]$ and Δt sufficiently small.

It can be shown that the weak order of convergence for EM method is $\gamma = 1$ [11].

6.2 Milstein’s Higher Order Method

It was mentioned above that EM has strong order of convergence $\gamma = \frac{1}{2}$ whereas the underlying deterministic Euler method converges with classical order 1. It is possible to raise the strong order of EM to 1 by adding a correction to the stochastic increment, resulting in Milstein’s method. The correction arises because the traditional Taylor expansion must be modified in the case of Itô calculus. Truncating the Itô–Taylor expansion at an appropriate point produces Milstein’s method for the SDE 6.1.1:

$$\begin{aligned} X_j &= X_{j-1} + f(X_{j-1}) + g(X_{j-1})(W(\tau_j) - W(\tau_{j-1}))\Delta t \\ &\quad + \frac{1}{2}g(X_{j-1})g'(X_{j-1})((W(\tau_j) - W(\tau_{j-1}))^2 - \Delta t), \\ j &= 1, 2, \dots, L \end{aligned} \quad (6.2.1)$$

6.3 Stochastic Runge-Kutta methods

High order numerical schemes can be further constructed by truncating Itô–Taylor expansion the same way it is done in EM and Milstein methods described above. But the computational cost can be high due to the proliferation of elementary derivatives. In order to obtain derivative-free methods, the extension of classical Runge-Kutta methods to stochastic differential equations has been introduced, see [6, 22].

Consider a scalar autonomous SDE

$$\begin{aligned} dy(t) &= a(t, y(t))dt + b(t, y(t))dW_t \quad t_0 \leq t \leq T \\ y(t_0) &= y_0 \end{aligned} \tag{6.3.1}$$

Let us introduce an equidistant discretization $\{t_0, \dots, t_L\}$ of the time interval $[0, T]$ with stepwise $\Delta = T/L$. The general form of an s -stage stochastic Runge-Kutta scheme for the solution of 6.3.1 is given by:

$$\left\{ \begin{array}{l} Y_i = y_n + \sum_{j=1}^s Z_{ij}^{(0)} a(t_n + \mu_j \Delta, Y_j) + \sum_{j=1}^s Z_{ij}^{(1)} b(t_n + \mu_j \Delta, Y_j) \\ i = 1, \dots, s \\ y_{n+1} = y_n + \sum_{j=1}^s z_j^{(0)} a(t_n + \mu_j \Delta, Y_j) + \sum_{j=1}^s z_j^{(1)} b(t_n + \mu_j \Delta, Y_j) \end{array} \right. \tag{6.3.2}$$

or, equivalently,

$$\left\{ \begin{array}{l} K_i^n = a(t_n + \mu_j \Delta, y_n + \Delta \sum_{j=1}^s \lambda_{ij} K_j^n + \Delta W \sum_{j=1}^s \gamma_{ij} \overline{K}_j^n) \quad i = 1, \dots, s \\ \overline{K}_i^n = b(t_n + \mu_j \Delta, y_n + \Delta \sum_{j=1}^s \lambda_{ij} K_j^n + \Delta W \sum_{j=1}^s \gamma_{ij} \overline{K}_j^n) \quad i = 1, \dots, s \\ y_{n+1} = y_n + \sum_{i=2}^s \alpha_i K_i^n + \sum_{i=1}^s \beta_i \overline{K}_i^n \end{array} \right. \tag{6.3.3}$$

with $\mu_1 = 0$. $Z^{(1)}$ and $z^{(1)}$ are respectively arbitrary matrix and vector whose elements are random variables and $Z^{(0)}$ and $z^{(0)}$ are respectively the parameter matrix and vector associated with the deterministic components.

If both $Z^{(0)}$ and $Z^{(1)}$ are strictly lower triangular, then 6.3.2 is said to be explicit, otherwise it is implicit.

The above scheme 6.3.2 or 6.3.3 can be represented in the tableau form

$$\frac{\mu \mid Z^{(0)} \mid Z^{(1)}}{z^{(0)T} \mid z^{(1)T}}$$

A specific case is

$$\left\{ \begin{array}{l} Y_i = y_n + \Delta \sum_{j=1}^s \lambda_{ij} a(t_n + \mu_j \Delta, Y_j) + \Delta W \sum_{j=1}^s \gamma_{ij} b(t_n + \mu_j \Delta, Y_j) \\ i = 1, \dots, s \\ y_{n+1} = y_n + \Delta \sum_{j=1}^s \alpha_j a(t_n + \mu_j \Delta, Y_j) + \Delta W \sum_{j=1}^s \beta_j b(t_n + \mu_j \Delta, Y_j) \end{array} \right. \quad (6.3.4)$$

and its tableau form for $s = 3$ is

$$\frac{\begin{array}{c} 0 \\ \mu_2 \\ \mu_3 \end{array} \mid \begin{array}{ccc} 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & \lambda_{32} & 0 \end{array} \mid \begin{array}{ccc} 0 & 0 & 0 \\ \gamma_{21} & 0 & 0 \\ \gamma_{31} & \gamma_{32} & 0 \end{array}}{\begin{array}{ccc} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{array}}$$

Analogously with the deterministic case, the technique for obtaining the order conditions consists in matching the truncated Runge-Kutta scheme with the stochastic Taylor series expansion of the exact solution over one step assuming exact initial values [22].

7 Kalman Filters

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discretedata linear filtering problem. Since that time, due

in large part to advances in digital computing, the Kalman filter has been the subject of extensive research and application, particularly in the area of mathematical finance, see for example [3, 9].

In this section we firstly describe the extended Kalman filter (EKF), then a promising approximation to EKF, called the Variational Kalman Filter (VKF) is reviewed, and finally we consider the hybrid method that combines the ensemble filtering approach with the VKF, prominent as the Variational Ensemble Kalman Filter (VEnKF).

7.1 Extended Kalman Filter

Many estimation problems that are of practical interest are nonlinear but "smooth". That is, the functional dependences of the measurement or state dynamics on the system state are nonlinear, but almost linear for small perturbations in the values of the state variables. For these cases the *Extended Kalman filter* can be applied. In this extension of the standard Kalman filter linearization techniques are employed to get linear approximations of the nonlinear system dynamics and measurement operators.

Suppose that the state $x \in \mathbb{R}^d$ of a discrete-time controlled process is governed by the nonlinear stochastic difference equation

$$\mathbf{x}_k = \mathcal{M}(\mathbf{x}_{k-1}) + \varepsilon_k^p \quad (7.1.1)$$

with a measurement $\mathbf{y} \in \mathbb{R}^m$ that is

$$\mathbf{y}_k = \mathcal{K}(\mathbf{x}_k) + \varepsilon_k^o \quad (7.1.2)$$

The random variables ε_k^p and ε_k^o represent the process and measurement noise (respectively). They are assumed to be independent (of each other), white, and with normal probability distributions

$$\begin{aligned} \varepsilon_k^p &\sim N(0, \mathbf{C}_{\varepsilon_k^p}) \\ \varepsilon_k^o &\sim N(0, \mathbf{C}_{\varepsilon_k^o}) \end{aligned} \quad (7.1.3)$$

where $\mathbf{C}_{\varepsilon_k^p}$ and $\mathbf{C}_{\varepsilon_k^o}$ are *process noise covariance* and *measurement noise covariance* matrices, correspondingly.

In the above models $\mathcal{M} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the *model evolution operator* that relates the state at the previous time step $k - 1$ to the state at the current step k , in the absence of either a driving function or process noise, and $\mathcal{K} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ is the *observation operator* that relates state to the measurement \mathbf{y}_k . The Extended Kalman filter algorithm for estimating states and their error covariances can be described as follows.

The Extended Kalman Filter algorithm

1. Compute the predicted state estimate and covariance:
 - (a) Compute the prior state $\mathbf{x}_k^p = \mathcal{M}(\mathbf{x}_{k-1}^{est})$
 - (b) Compute the prior covariance $\mathbf{y}_k^p = \mathcal{K}(\mathbf{x}_k^p)$
 - (c) Compute linearized model operator $\mathbf{M}_k = \partial (\mathcal{M}(\mathbf{x}_{k-1}^{est})) \partial x$
 - (d) Compute linearized measurement operator $\mathbf{K}_k = \partial (\mathcal{K}(\mathbf{x}_k^p)) \partial x$
 - (e) Compute the prior covariance $\mathbf{C}_k^p = \mathbf{M}_k \mathbf{C}_{est_{k-1}}^T \mathbf{M}_k^T + \mathbf{C}_{\varepsilon_k^p}$
2. Combine the prior with observations:
 - (a) Compute the Kalman gain $\mathbf{G}_k = \mathbf{C}_k^p \mathbf{K}_k^T (\mathbf{K}_k \mathbf{C}_k^p \mathbf{K}_k^T + \mathbf{C}_{\varepsilon_k^o})^{-1}$
 - (b) Compute the state estimate $\mathbf{x}_k^{est} = \mathbf{x}_k^p + \mathbf{G}_k (\mathbf{y}_k - \mathbf{K}_k \mathbf{x}_k^p)$
 - (c) Compute the covariance estimate $\mathbf{C}_k^{est} = \mathbf{C}_k^p - \mathbf{G}_k \mathbf{K}_k \mathbf{C}_k^p$
3. Set $k \rightarrow k + 1$ and go to step 1

Since the standard formulation of the Kalman filter (KF) and extended Kalman filter (EKF) require the storage and multiplication of $d \times d$ matrices and the inversion of $m \times m$ matrices, for large-scale problems these methods suffer from the high CPU and memory requirements. However,

these problems can be overcome by using a low-storage approximation, called Variational Kalman Filter (VKF).

7.2 Variational Kalman Filter

In variational formulation of Kalman filter the large matrices of linearized evolution and observation operators were replaced with the tangent linear and adjoint codes, that produce differentiation at the 'code level'. These codes are used to propagate covariance in time. The state estimation itself is viewed as an optimization problem, where the quadratic function

$$l(\mathbf{x}|y_k) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_k^p)^T (\mathbf{C}_k^p)^{-1} (\mathbf{x} - \mathbf{x}_k^p) + \frac{1}{2}(\mathbf{y} - \mathcal{K}(\mathbf{x}))^T (\mathbf{C}_{\varepsilon_k^o})^{-1} (\mathbf{y} - \mathcal{K}(\mathbf{x})) \quad (7.2.1)$$

is minimized with respect to \mathbf{x} . In the VKF method, introduced in [2] the minimization is done with the Limited Memory Broyden-Fletcher-Goldfarb-Shanno (LBFSGS) optimization method, that produces both the state estimate and the covariance approximation. The inverse of the prior covariance is also approximated via solving an auxiliary optimization problem, given as

$$\arg \min_{\mathbf{u}} \frac{1}{2} \mathbf{u}^T \mathbf{C}_k^p \mathbf{u} \quad (7.2.2)$$

Supposing that the linearization \mathbf{M}_k of $\mathcal{M}(x_{k-1}^{est})$ is available, the algorithmic representation for the nonlinear VKF method can be written as follows.

The Variational Kalman Filter algorithm

1. Compute the predicted state estimate and covariance:
 - (a) Compute the prior state $\mathbf{x}_k^p = \mathcal{M}(\mathbf{x}_{k-1}^{est})$
 - (b) Define $\mathbf{C}_k^p = \mathbf{M}_k \mathbf{C}_{k-1}^{est} \mathbf{M}_k^T + \mathbf{C}_{\varepsilon_k^p}$
 - (c) Apply LBFSGS to 7.2.2 to get an approximation of $(\mathbf{C}_k^p)^{-1}$

2. Combine the prior with observations:

(a) Minimize expression 7.2.1 using LBFGS to get the state estimate

$$x_k^{est} \text{ and covariance estimate } \mathbf{C}_k^{est}.$$

3. Set $k \rightarrow k + 1$ and go to step 1

7.3 Variational Ensemble Kalman Filter

In the Variational Ensemble Kalman Filter (VEnKF) the VKF ideas are employed in ensemble filtering context. As in VKF, the state estimation in VEnKF is based on minimizing the cost function in equation 7.2.1. The prior covariance required in the cost function is defined as

$$\begin{aligned} \mathbf{C}_k^p &= \text{Cov}(\mathcal{M}(\mathbf{x}_{k-1}^{est}) + \varepsilon_k^p) = \text{Cov}(\mathcal{M}(\mathbf{x}_{k-1}^{est})) + \text{Cov}(\varepsilon_k^p) \\ &\approx \mathbf{X}_k \mathbf{X}_k^T + \mathbf{C}_{\varepsilon_k^p} \end{aligned} \quad (7.3.1)$$

In the above formula it is assumed that the model error and model response are uncorrelated. In VEnKF the sample covariance is calculated as

$$\mathbf{X}_k = ((\mathbf{s}_{k,1} - x_k^p), (\mathbf{s}_{k,2} - x_k^p), \dots, (\mathbf{s}_{k,N} - x_k^p)) / \sqrt{N - 1} \quad (7.3.2)$$

where the state estimate evolved from the previous time is used as the expectation value ($x_k^p = \mathcal{M}(x_{k-1}^{est})$), $\mathbf{s}_k = (\mathbf{s}_{k,1}, \mathbf{s}_{k,2}, \dots, \mathbf{s}_{k,N})$ denotes the ensemble and $s_{k,i} = \mathcal{M}(s_{k-1,i})$. The inverse prior covariance matrix $\mathbf{C}_k^p = \mathbf{X}_k \mathbf{X}_k^T + \mathbf{C}_{\varepsilon_k^p}$ can be acquired either by applying the LBFGS to the artificial optimization problem

$$\arg \min_{\mathbf{u}} \mathbf{u}^T (\mathbf{X}_k \mathbf{X}_k^T + \mathbf{C}_{\varepsilon_k^p}) \mathbf{u} \quad (7.3.3)$$

or using the Sherman-Morrison-Woodbury (SMW) matrix inversion formula:

$$\begin{aligned} (\mathbf{C}_k^p)^{-1} &= (\mathbf{X}_k \mathbf{X}_k^T + \mathbf{C}_{\varepsilon_k^p})^{-1} \\ &= \mathbf{C}_{\varepsilon_k^p}^{-1} - \mathbf{C}_{\varepsilon_k^p}^{-1} \mathbf{X}_k (\mathbf{I} + \mathbf{X}_k^T \mathbf{C}_{\varepsilon_k^p}^{-1} \mathbf{X}_k) \mathbf{X}_k^T \mathbf{C}_{\varepsilon_k^p}^{-1} \end{aligned} \quad (7.3.4)$$

The algorithmic description of the VEnKF method is given below, see [24].

The VEnKF algorithm

1. Move the ensemble forward and build the prior:
 - (a) Compute the prior center point $\mathbf{x}_k^p = \mathcal{M}(\mathbf{x}_{k-1}^{est})$
 - (b) Compute prior ensemble $s_{k,i} = \mathcal{M}(s_{k-1,i}), i = 1, \dots, N$
 - (c) Define $(\mathbf{C}_k^p)^{-1}$ using SMW formula 7.3.4 or applying the LBFGS to to 7.3.3.
2. Calculate the posterior estimate and generate the new ensemble:
 - (a) Apply LBFGS to minimize 7.2.1 to get the state estimate x_k^{est} and the covariance estimate \mathbf{C}_k^{est} .
 - (b) Sample new ensemble $s_k^{est} \sim N(\mathbf{x}_k^{est}, \mathbf{C}_k^{est})$
3. Set $k \rightarrow k + 1$ and go to step 1

In order to sample new ensemble of state vectors from $N(\mathbf{x}_k^{est}, \mathbf{C}_k^{est})$ one can use the LBFGS representation for \mathbf{C}_k^{est} :

$$\mathbf{C}_k^{est} = \mathbf{B}_0 \mathbf{B}_0^T + \sum_{i=1}^n \mathbf{b}_i \mathbf{b}_i^T \quad (7.3.5)$$

where \mathbf{B}_0 is $d \times d$ matrix and \mathbf{b}_i are $d \times 1$ vectors. From this representation a zero mean random vector $\mathbf{r} \sim N(\mathbf{0}, \mathbf{C}_k^{est})$ can be produced by calculating

$$\mathbf{r} = \mathbf{B}_0 \mathbf{z} + \sum_{i=1}^n \omega_i \mathbf{b}_i \quad (7.3.6)$$

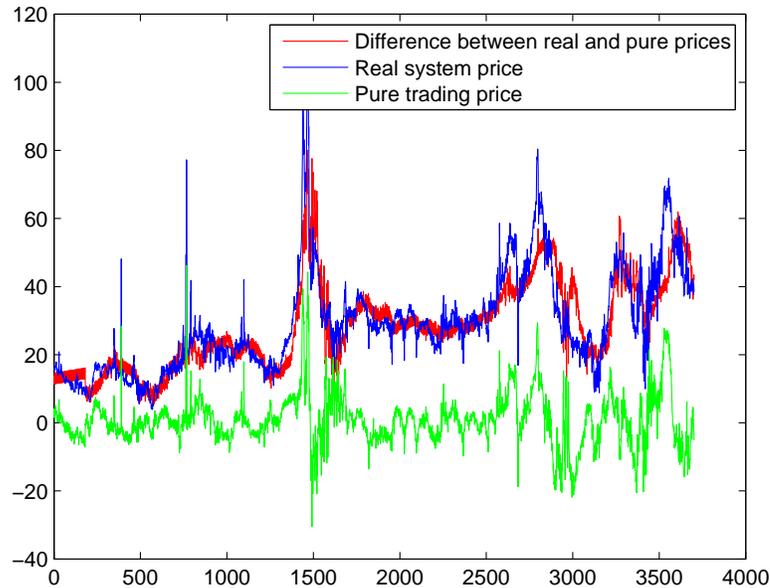
where $\mathbf{z} \sim N(\mathbf{0}, \mathbf{1})$ and $\omega_i \sim N(0, 1)$, \mathbf{B}_0 and \mathbf{b}_i can be constructed from LBFGS inverse Hessian approximation, see [24].

8 A fluid mechanic model of Nordpool spot prices

In this section we describe the model of Nordpool spot prices based on the Burgers' equation with an external forcing that consists on three terms.

The first forcing term p produces the low-speed motions that experience a mean-reverting behavior. This component was introduced by analogy with the governing equations for atmospheric motions and operates similar to the geostrophic force that results in baroclinic Rossby waves. The second forcing component, described by Brownian motion term dB_t , implies the normally distributed random fluctuations about the general path. Finally, $f(t)$ contains the difference between the real and the pure trading prices (see Figure 5), that is, it represents the influence of the linear regression components that are responsible for seasonal behavior and pursuing the main trend.

Figure 5: Real system price (blue) pure trading price (green) and their residual (red)



Thereby, we consider the following initial-boundary value problem for Stochastic Partial differential equation:

$$\begin{aligned} -du(x, t) + pdt + \alpha u(x, t)u_x(x, t)dt \\ + \vartheta u_{xx}(x, t)dt = f(t)dt + dB_t, \quad 0 \leq x \leq 1 \end{aligned} \quad (8.1)$$

$$u(x, 0) = u_0(x) \quad (8.2)$$

$$u(0, t) = u(1, t), \quad (8.3)$$

where $p = u - \mathbf{MA}_{180}$ contains the difference between the current system price and the system price six month moving average.

Applying first-order operator splitting approach (see [23]), we decouple the convection-reaction and diffusion operators at the each time interval $[t_k, t_{k+1}]$:

Convection-reaction subproblem

$$\begin{aligned} -du + pdt + \alpha uu_x dt = f(t)dt + dB_t, \quad 0 \leq x \leq 1 \\ u(x, 0) = u_0(x) \end{aligned} \quad (8.4)$$

$$u(0, t) = u(1, t)$$

A solution of the convection-reaction subproblem $u^{(*)}$ defines the initial conditions for the diffusion subproblem:

Diffusion subproblem

$$\begin{aligned} -du + \vartheta u_{xx} dt = 0, \quad 0 \leq x \leq 1 \\ u(x, 0) = u^{(*)}(x) \end{aligned} \quad (8.5)$$

$$u(0, t) = u(1, t)$$

Employing the semi-Lagrangian formulation, the problem 8.4 can be rewritten as follows:

$$\begin{aligned} Du = pdt - f(t)dt - dB_t \\ u(x, 0) = u_0(x) \end{aligned} \quad (8.6)$$

$$u(0, t) = u(1, t)$$

where Du denotes the total derivative of u along the trajectory

$$Du = du + \frac{dx}{dt} \frac{\partial u}{\partial x} dt = du - \alpha u \frac{\partial u}{\partial x} dt \quad (8.7)$$

The last equality 8.7 implies the characteristic equation

$$x'(t) = -\alpha u(x(t), t) \quad (8.8)$$

According to equation 8.8, we firstly advect the grid points to obtain a solution of the problem 8.6 via the method of characteristics:

$$\tilde{\mathbf{x}} = \mathbf{x} - \alpha \mathbf{u}^k \tau, \quad (8.9)$$

where \mathbf{x} denotes the fixed set of grid points, \mathbf{u}^k stands for the vector of solution from the previous time step, τ denotes the time stepping. Taking into account for the periodical boundary conditions, we transfer the advected points $\tilde{\mathbf{x}}$ onto the initial spacial domain $[0, 1]$:

$$\mathbf{i} = \lfloor \frac{\tilde{\mathbf{x}}}{h} \rfloor \quad (8.10)$$

$$\mathbf{j} = \text{mod}(\mathbf{i}, N_x - 1) \quad (8.11)$$

where h stands for the uniform distance between the spacial points, N_x denotes the number of points and \mathbf{j} contains the numbers of the nearest left grid points for the counterparts of advected points.

$$\mathbf{d} = \tilde{\mathbf{x}} - \mathbf{i}h \quad (8.12)$$

Here \mathbf{d} determines the distances between the projected points and the corresponding closest left points of the fixed grid.

$$\tilde{\tilde{\mathbf{x}}} = \mathbf{j}h + \mathbf{d}, \quad (8.13)$$

Consequently, we apply the following Euler-Maruyama discretization to solve the Cauchy problem 8.6:

$$\hat{\mathbf{u}}^k = \tilde{\tilde{\mathbf{u}}}^k + \mathbf{p}\tau - \mathbf{f}^k\tau - \mathbf{dB}^k \quad (8.14)$$

where

- $\mathbf{dB}^k \sim \sqrt{\tau}N(\mathbf{0}, \mathbf{I}_{N_x \times N_x})$ is a multivariate Wiener gain;

- $\mathbf{f}^k = f(t_k)\mathbf{I}_{N_x \times 1}$;
- $\tilde{\mathbf{u}}^k$ denotes the initial condition \mathbf{u}^k , interpolated onto the set $\tilde{\mathbf{x}}$;
- $\mathbf{p} = \tilde{\mathbf{u}}^k - \text{MA}_{180}\mathbf{I}_{N_x \times 1}$.

Finally, we approximate the solution of the diffusion subproblem 8.5:

$$\mathbf{u}_i^{k+1} = \hat{\mathbf{u}}_i^k + \vartheta \frac{\hat{\mathbf{u}}_{i+1}^k - 2\hat{\mathbf{u}}_i^k + \hat{\mathbf{u}}_{i-1}^k}{h^2}, \quad i = 2, \dots, N_x - 1 \quad (8.15)$$

$$\mathbf{u}_1^{k+1} = \hat{\mathbf{u}}_1^k + \vartheta \frac{\hat{\mathbf{u}}_2^k - 2\hat{\mathbf{u}}_1^k + \hat{\mathbf{u}}_{N_x}^k}{h^2}, \quad (8.16)$$

$$\mathbf{u}_{N_x}^{k+1} = \hat{\mathbf{u}}_{N_x}^k + \vartheta \frac{\hat{\mathbf{u}}_1^k - 2\hat{\mathbf{u}}_{N_x}^k + \hat{\mathbf{u}}_{N_x-1}^k}{h^2}. \quad (8.17)$$

9 Applying Kalman filtering to forecasting spot prices

In order to adjust a prediction for Nordpool system price, any modification of the standart Kalman filter which is suitable for nonlinear dynamic systems (i.e., EKF, VKF or VEnKF) can be possibly exploited. However, in this work we are concerned on the VEnKF application.

Out of the N_x model states, measurement is taken from the only one state. At the every time step we build the histogram out of the estimated data and determine a mode. In case of a multimodal distribution we consider the first mode value. Thereby, the observation operator is defined as $\mathcal{K}x = \mathbf{K}x$, where

$$[\mathbf{K}]_{rp} = \begin{cases} 1, & r = p = m, \\ 0, & \text{otherwise.} \end{cases} \quad (9.1)$$

Here m is a number of an arbitrary state, where the mode value was observed in the previous estimated data.

Since for application of this hybrid filter an ensemble of particles has to be produced, we generate a new ensemble before performing each successive

estimation. Here we take into account the fact, that the traders observe one another within the small microgroup. It is also assumed, that the observable neighbourhood composes 5% of the whole market traders. Hence, we firstly build the histogram with 20 bins out of estimated data and determine the mode value η . Thereafter, we produce an ensemble $\mathbf{s} \sim N(\boldsymbol{\eta}, \boldsymbol{\Sigma})$ from multivariate normal distribution, where $\boldsymbol{\eta} = \eta \cdot \mathbf{I}_{N_x \times 1}$, $\boldsymbol{\Sigma}$ is a diagonal matrix, where the diagonal elements contain the squared difference between the mode η and the value of the closest neighbouring bin.

We run experiments with varying ensemble sizes N , model and observation error covariances are assumed to be diagonal and constant; they are given by $\mathbf{C}_{\varepsilon_k^p} = (0.05\sigma_{clim})^2 \mathbf{I}_{N_x \times N_x}$ and $\mathbf{C}_{\varepsilon_k^o} = (0.15\sigma_{clim})^2 \mathbf{I}_{N_x \times N_x}$, correspondingly, where $\sigma_{clim} = 3.641$ (standart deviation used in climatological simulations), see [24].

As the initial guesses in filtering we use $\mathbf{x}_0^{\text{est}} \sim N(\mu, \sigma^2 \mathbf{I}_{N_x \times 1})$, where μ and σ are fixed arbitrary constants, and $\mathbf{C}_0^{\text{est}} = \mathbf{I}_{N_x \times N_x}$. Estimations are produced taking either real system or pure trading prices as an observation.

10 Results

After applying the VEnKF to the specified model, we plotted the estimated mode together with observed data, see Figures 6, 7, 8. On purpose to observe the convergence properties, we calculate the root mean square error (rms), written as

$$[rmse]_k = |x_k^{mo} - x_k^{true}| \quad (10.1)$$

where x_k^{mo} is estimated ensemble mode, x_k^{true} is the observed price, see Figure 10, 11, 12. In case of pure trading series rmse error decreases to almost zero when the size of the ensemble is large enough (e.g., $N = 40$), see Figure 12.

From relative error plot (see Figure 9) one can conclude that the convergence

for the real system price is rather slow and it breaks down in case of spike derivation. We have also plotted rmse error averaged over time, see Figure 13. One can observe directly proportional relationship between the size of the ensemble and the forecast skill.

In order to compare the model performance for fixated rate of advection ($\alpha = 5/3$), we generate 20, 40 and 60 days ahead forecasts for different diffusion coefficients ϑ , starting from the every n -th estimate ($n = 20, 40, 60$, correspondingly, see Figures 14, 16, 18). From obtained forecasts we firstly observe the model reaction upon the spikes. Thereof, we state that in the presence of diffusion the spikes are systematically followed by spurious subsequently decreasing oscillations. In addition, it can be noticed that the stronger is the diffusion, the larger is the amplitude of the oscillations.

In the absence of diffusion the model reacts upon monotonic increase or decrease, producing spurious spikes. The same phenomena is observed in the presence of diffusion, so that the bigger is the diffusion coefficient, the stronger is oscillation.

Fixating the diffusion coefficient ($\theta = 0.08$), we investigate the dependence of model model behavior on the rate of advection. From the Figures 20, 22, 20 one can observe that the stronger is the advection, the sharper is the forecast deviation from the current level.

Figure 6: Ensemble mode (blue) and pure trading price (red) for ensemble size $N = 10$, $\alpha = 5/3$, $\theta = 0.08$

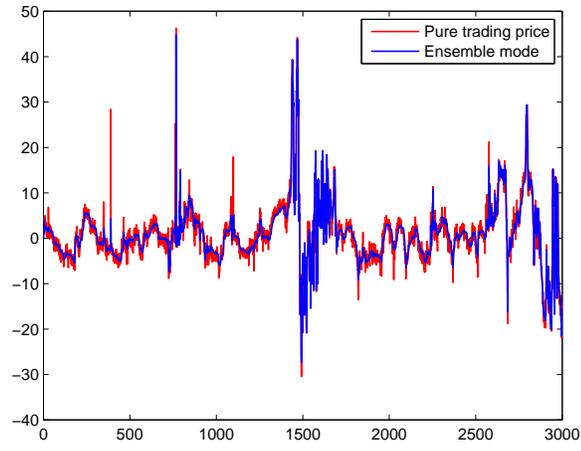


Figure 7: Ensemble mode (blue) and pure trading price (red) for ensemble size $N = 40$, $\alpha = 5/3$, $\theta = 0.08$

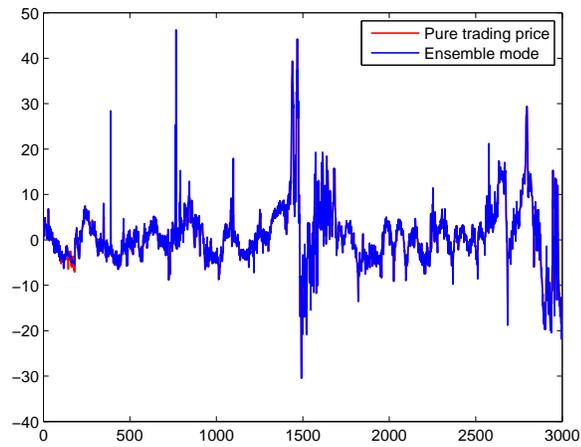


Figure 8: Ensemble mode (blue) and real system price (red) for ensemble size $N = 10$, $\alpha = 5/3$, $\theta = 0.08$

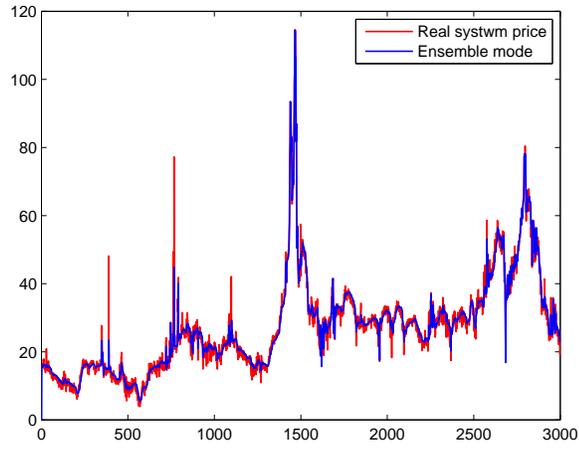


Figure 9: Relative error plot for real system price, $N = 10$, $\alpha = 5/3$, $\theta = 0.08$

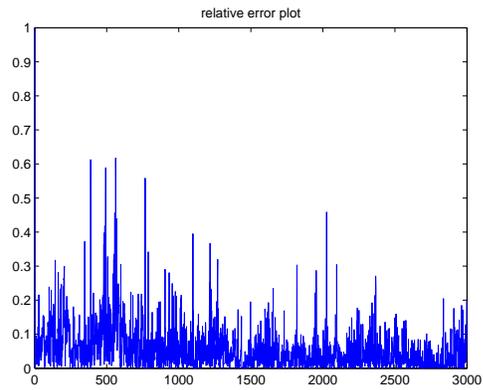


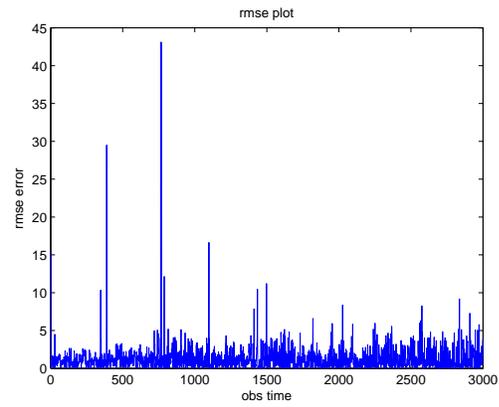
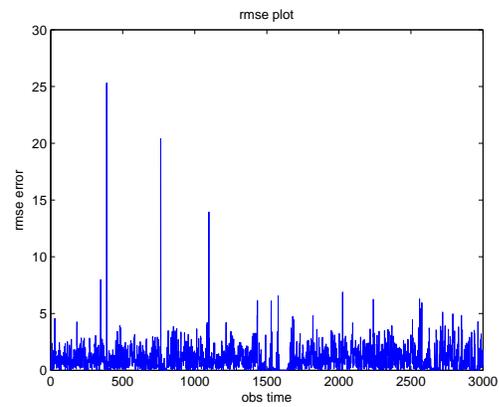
Figure 10: Rmse error for real system price, $N = 10$, $\alpha = 5/3$, $\theta = 0.08$ Figure 11: Rmse error for pure trading price for ensemble size $N = 10$, $\alpha = 5/3$, $\theta = 0.08$ 

Figure 12: Rmse error for pure trading price for ensemble size $N = 40$, $\alpha = 5/3, \theta = 0.08$

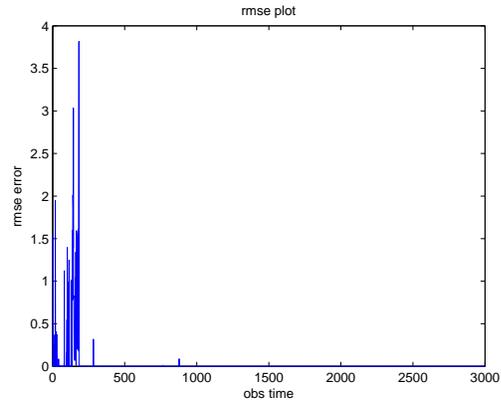


Figure 13: Forecast skills for Real system price for different ensemble size : $N=10$ (blue), $N=15$ (green), $N=30$ (yellow), $N=40$ (red) $\alpha = 5/3, \theta = 0.001$

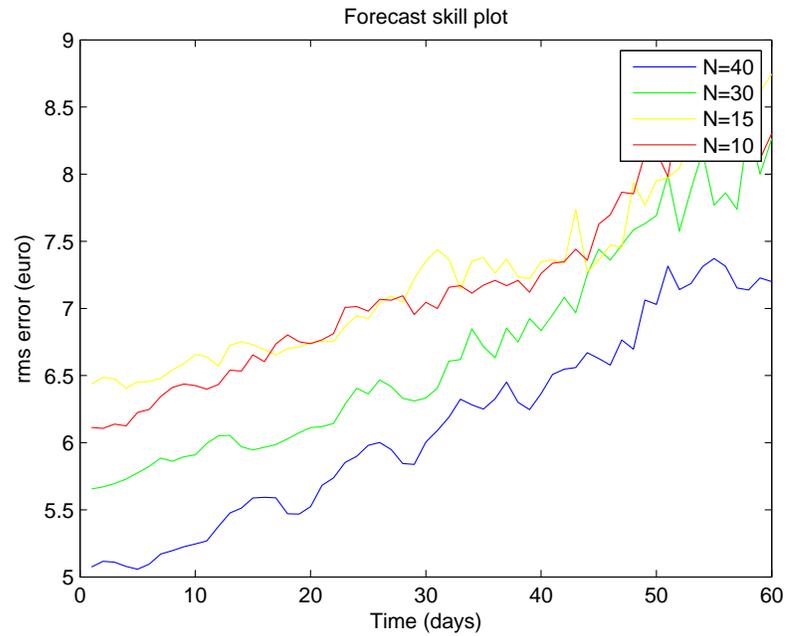
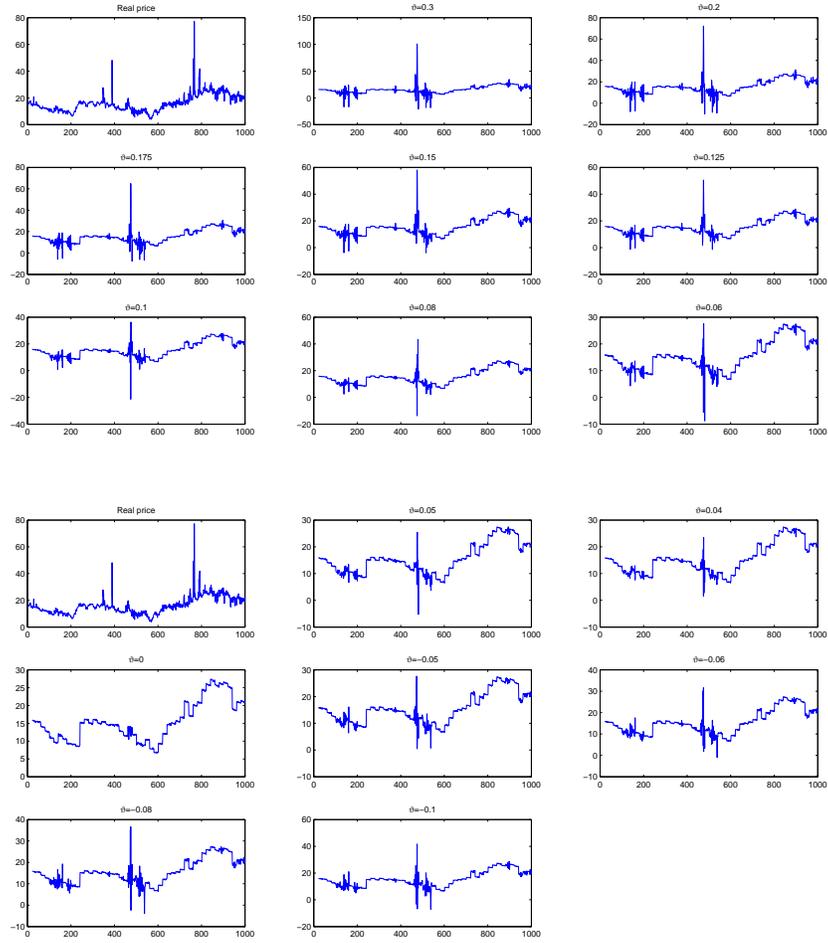


Figure 14: Real price and 20 days ahead forecasts, $\alpha = 5/3$



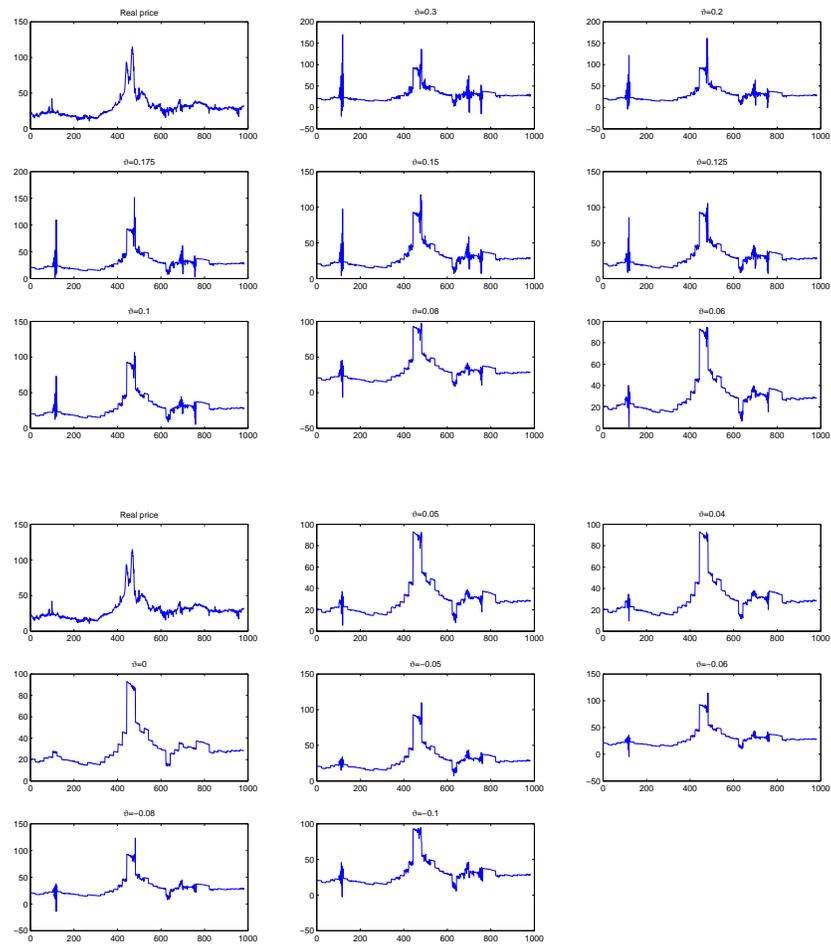
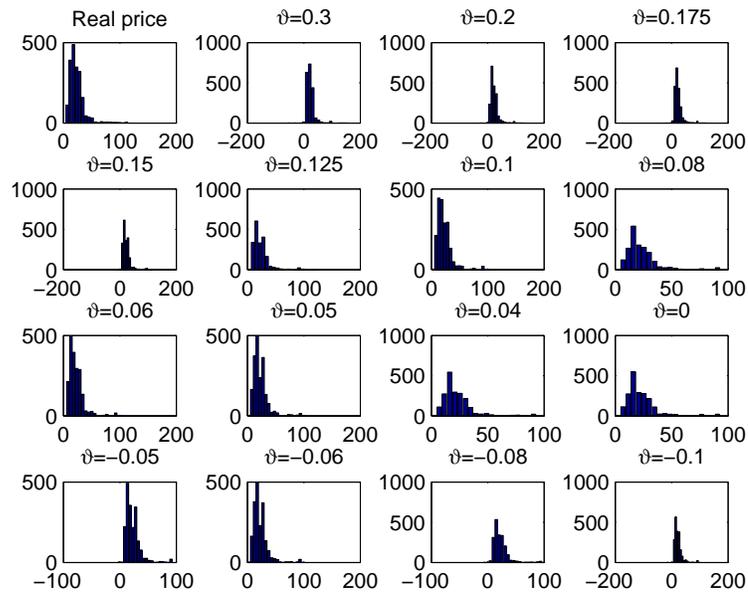
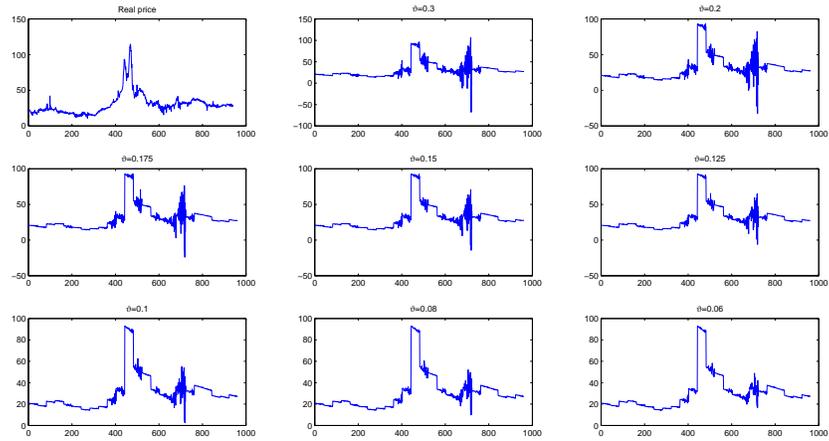


Figure 15: Distributions of real price and 20 days ahead forecasts, $\alpha = 5/3$ Figure 16: Real price and 40 days ahead forecasts, $\alpha = 5/3$ 

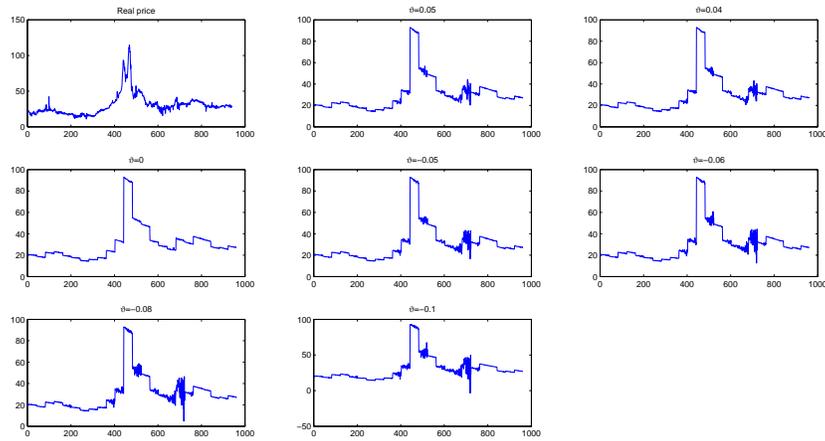


Figure 17: Distributions of real price and 40 days ahead forecasts, $\alpha = 5/3$

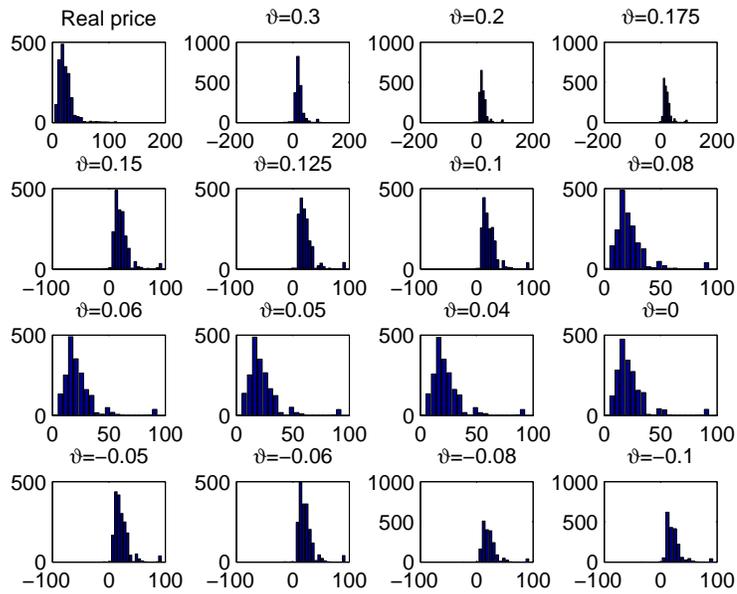


Figure 18: Real price and 60 days ahead forecasts, $\alpha = 5/3$

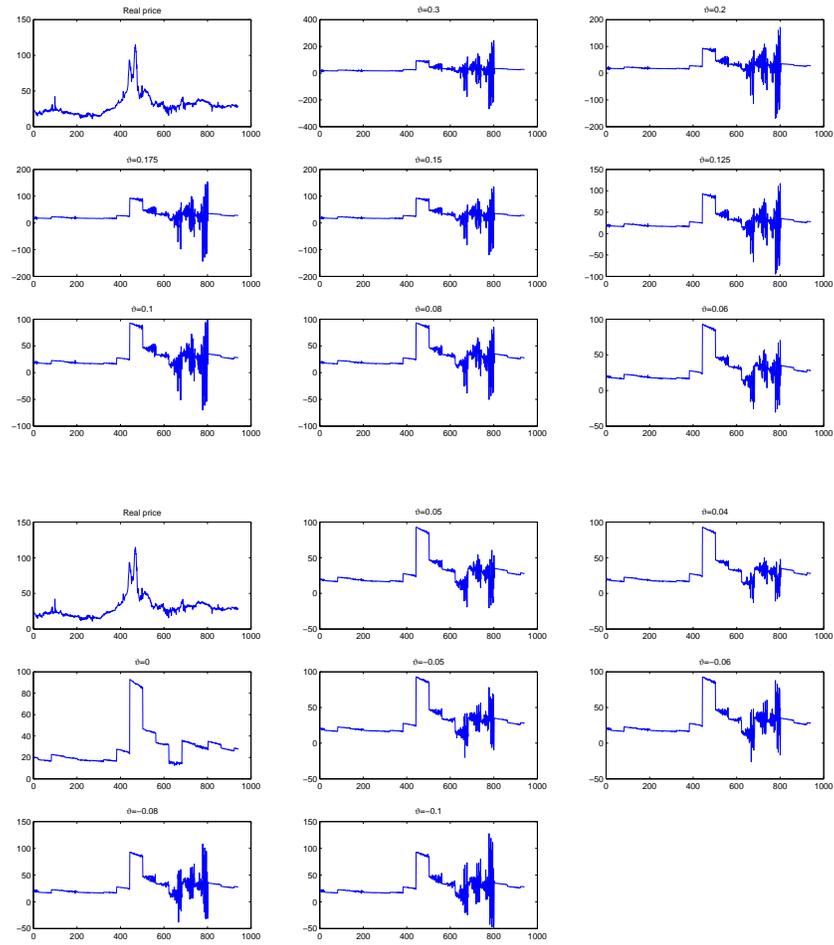


Figure 19: Distributions of real price and 60 days ahead forecasts, $\alpha = 5/3$

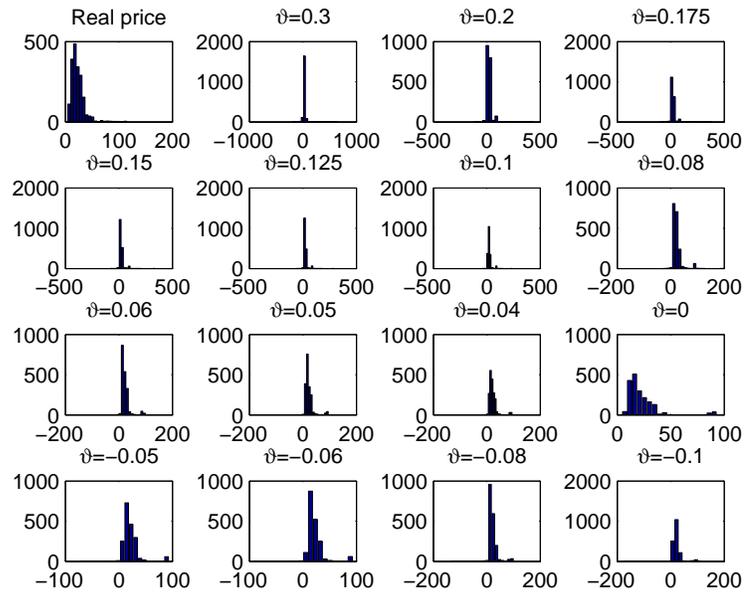
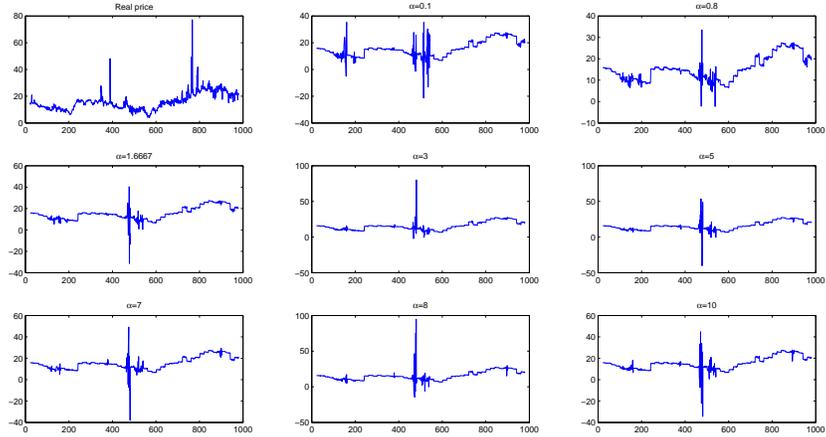


Figure 20: Real price and 20 days ahead forecasts, $\theta = 0.08$



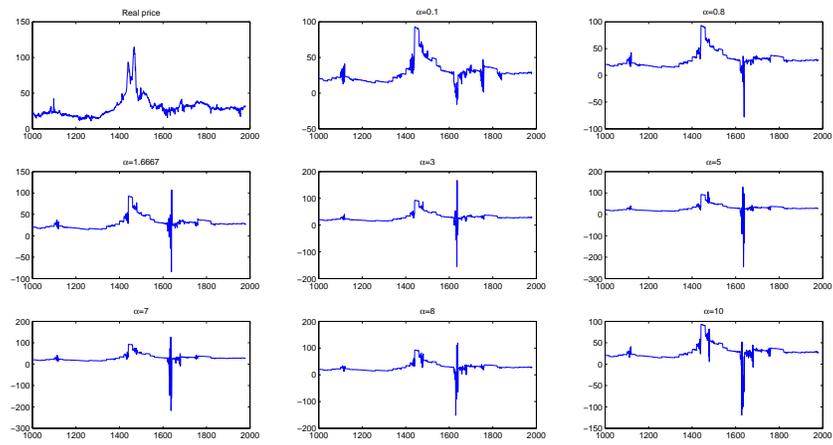


Figure 21: Distributions of real price and 20 days ahead forecasts, $\theta = 0.08$

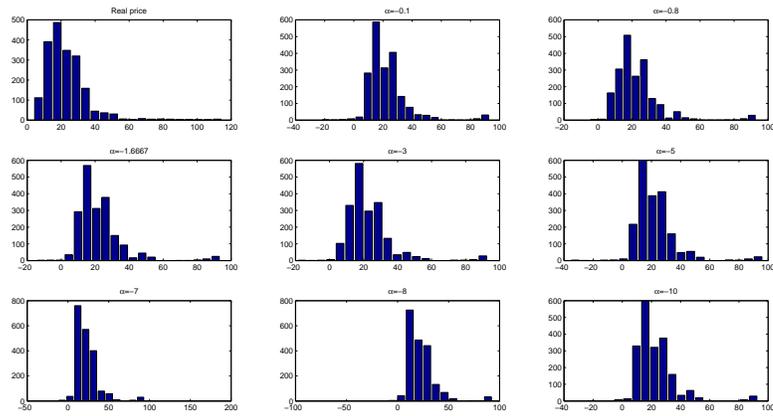


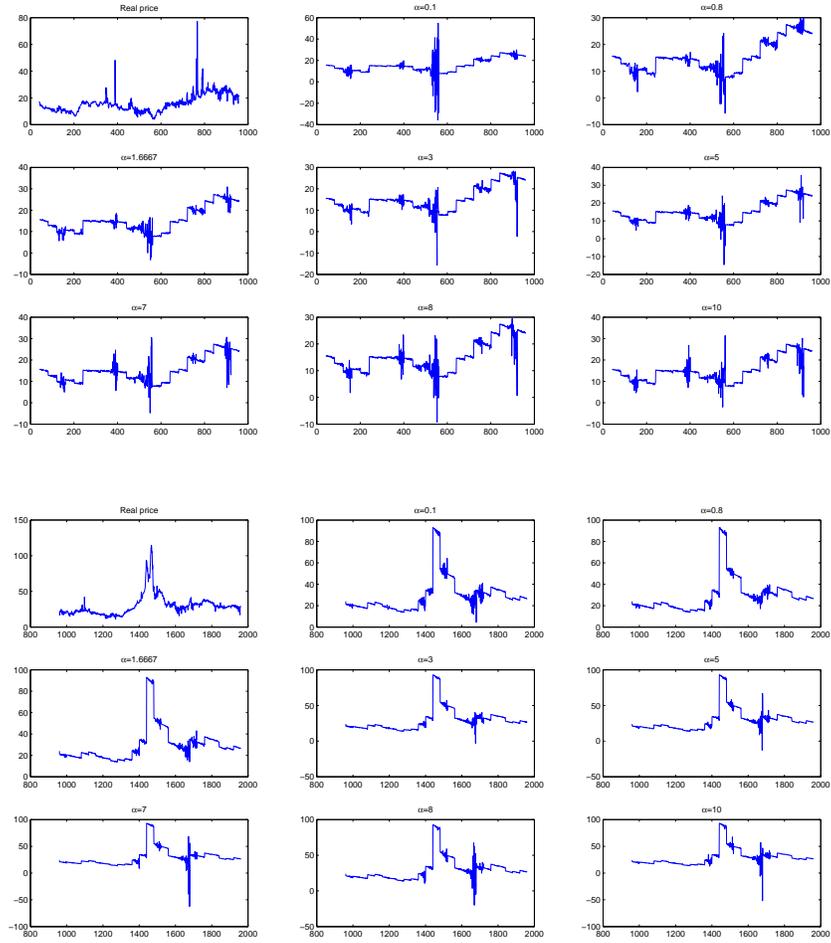
Figure 22: Real price and 40 days ahead forecasts, $\theta = 0.08$ 

Figure 23: Distributions of real price and 40 days ahead forecasts, $\theta = 0.08$

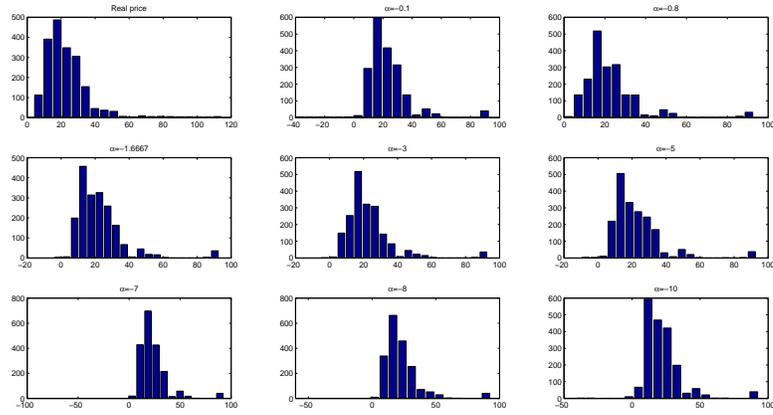
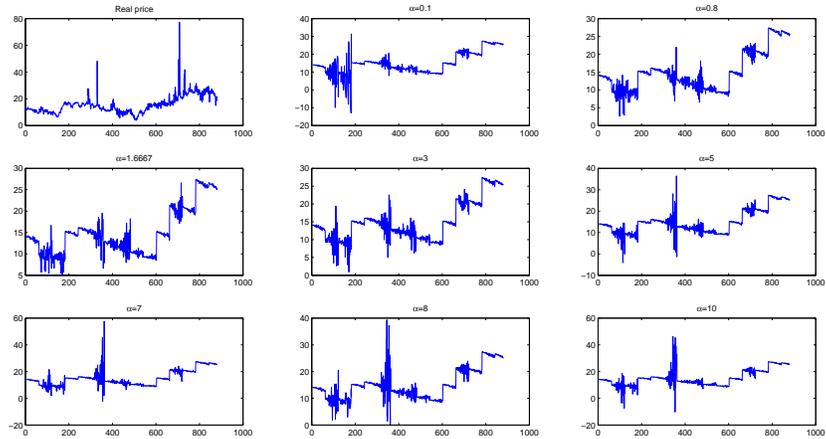


Figure 24: Real price and 60 days ahead forecasts, $\theta = 0.08$



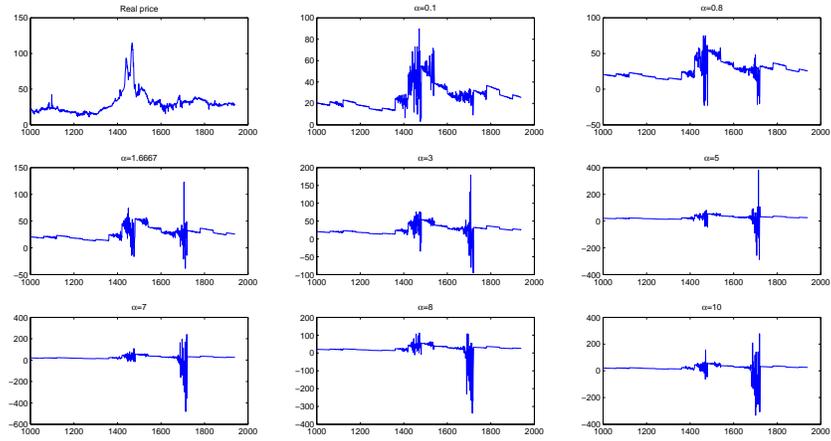


Figure 25: Distributions of real price and 60 days ahead forecasts, $\theta = 0.08$

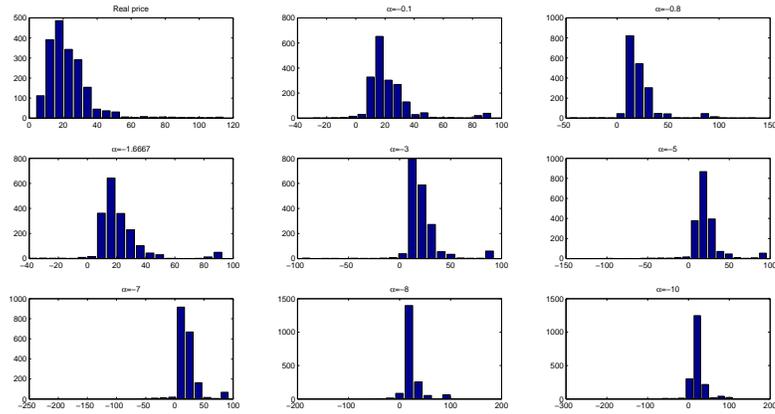


Figure 26: Distributions of pure trading price and 20 days ahead forecasts, $\alpha = 5/3$

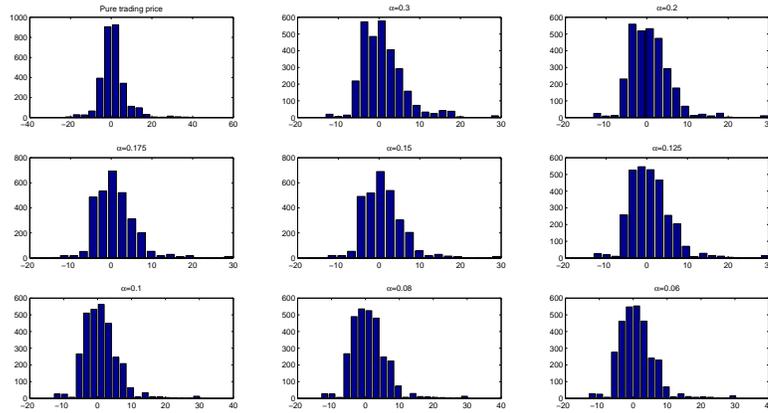
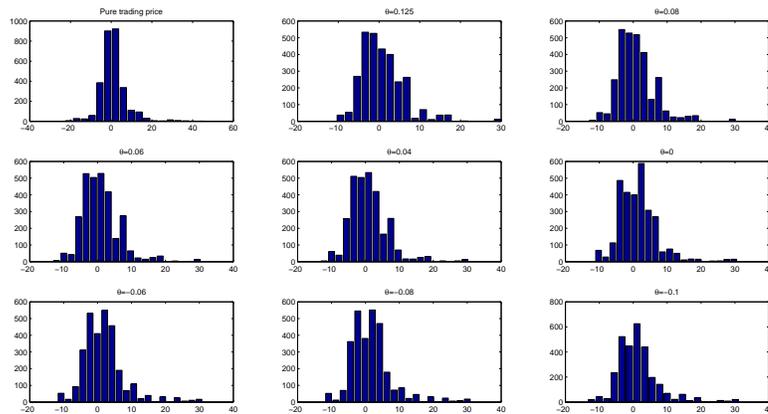


Figure 27: Distributions of pure trading price and 60 days ahead forecasts, $\alpha = 5/3$



11 Discussion

As we have observed, the specified model of Nord Pool spot prices considered in periodical spatial domain experiences a spurious volatility after spike

appearance, at the same time, due to influence of diffusion, large oscillations subsequently decay. We state that this "panic" reaction of the system is caused by the impetus contributed by the abrupt increase.

At the same time, similar strong volatility change is observed after period of long monotonic increase succeeded by decrease of the same character, that is taking place independently of diffusion rate.

In future work it is reasonable to consider the above-specified model in semi-infinite domain, imposing an appropriate boundary conditions. This choice is motivated by fundamental properties of the market, since in reality the traders observe the market tendencies locally and in the most cases they are not significantly influenced by the price changes outside their microgroup.

It might be also successful to concentrate upon the multidimensional models from fluid mechanics, such as Shallow Water Equations or any other system reduced from the Navier-Stocks equations.

References

- [1] Apte A., Auroux D., Ramaswamy M.: Variational data assimilation for discrete Burgers equation. In: *Electronic Journal of Differential Equations*, vol. 19, pp. 15-30, 2010.
- [2] Auvinen H., Bardsley J.M., Haario H. and Kauranne T.: The variational Kalman filter and an efficient implementation using limited memory BFGS. In: *Int. J. Numer. Meth. Fluids*, 2008. doi:10.1002/fld.2153, 2009.
- [3] Babbs S.H., Nowman K.B.: Kalman filtering of generalized Vasicek term structure models. In: *The Journal of Financial and Quantitative Analysis*, 1999.
- [4] Baya H., Buchasia C., Rykfors M., Saint-Aubain P. d., Vecchio I., Wangwe I., Yedeg E.: Influence of physical factors on electricity spot market price, Project work report, 23rd ECMI Modelling Week, Wrocław, Poland, 2009.
- [5] Bunn, D., Karakatsani, N.: Forecasting electricity prices, London Business School Working Paper, 2003.
- [6] Burrage, K., Platen, E.: Runge-Kutta methods for stochastic differential equations. In: *Ann. Numer. Math.* 1, pp. 63–78, 1994.
- [7] Deng, S.: Stochastic models of energy commodity prices and their applications: Mean reversion with jumps and spikes. Working paper. Georgia institute of Technology, 2000.
- [8] Ghashghaie S., Breymann W., and et al.: Turbulent cascades in foreign exchange markets. In: *Nature*, vol. 381: pp. 767-770, 1996.
- [9] Harvey A. C. : Forecasting, Structural Time Series and the Kalman Filter, Cambridge: Cambridge University, 1991.

- [10] Hayashi, F.: *Econometrics*, Princeton University Press, New Jersey, 2000.
- [11] Higham, D.J.: *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*. In: *SIAM Review*, vol. 43, no. 3, pp . 525–546, 2001
- [12] Huisman, R., Mahieu, R.: *Regime jumps in electricity prices*, *Energy Economics* 25, 425-434, 2003.
- [13] Jabłońska, M., Viljainen, S., Partanen, J., Kauranne, T.: *Regression Model-Based Analysis of the Impact of Emissions Trading on Electricity Spot Market Price Behavior*, *IEEE transactions on power systems*, vol. 10, no. 10, 2010.
- [14] Johnson, B., Barz, G.: *Selecting stochastic processes for modelling electricity prices*. In: *Energy Modelling and the Management of Uncertainty*, Risk Books, London, 1999.
- [15] Kaminski, V.: *The challenge of pricing and risk managing electricity derivatives*. In: *The US Power Market*, Barber, P. (ed.), Risk Books, London, 1997.
- [16] Karakatsani, N., Bunn, D.: *Modelling stochastic volatility in high-frequency spot electricity prices*, London Business School Working Paper, 2004.
- [17] Kirabo, B. N.: *Analysis of patterns in electricity spot market time series.*, Master's Thesis, Lappeenranta University of Technology, 2010.
- [18] LeVeque R., *Numerical Methods for Conservation Laws*, Lectures in Mathematics, Birkhäuser, Basel 1992.
- [19] Mantegna R.N., Palágyi Z., and Stanley H.E.: *Applications of statistical mechanics to finance*, *Physica A*, vol. 274, 1999. : p. 216-221.

- [20] Morale D., Capasso V., Oelschläger K.: An interacting particle system modelling aggregation behavior: from individuals to populations. In: *Journal of Mathematical Biology*, 2004.
- [21] Müller U.A., Dacorogna M.M., and et al.: Volatilities of different time resolutions - Analyzing the dynamics of market components. In: *J. Empirical Finance*, vol.4: p. 213-139, 1997.
- [22] Napoli, A.: Economical Runge-Kutta methods with weak second order for Stochastic Differential Equations. In: *Int. J. Contemp. Math. Sciences*, vol. 5, no. 24, pp. 1151 - 1160, 2010.
- [23] Smolianski A., Shipilova O., Haario H.: A Fast High-Resolution Algorithm for Linear Convection Problems: Particle Transport Method. In: *Int. J. Numer. Meth. Eng.*, vol. 70, pages 655-684, 2007.
- [24] Solonen A., Hakkarainen J., Auvinen H., Amour I., Haario H., Kau-ranne T.: Variational Ensemble Kalman Filtering Using Limited Mem-ory BFGS. In: *Int. J. Numer. Meth. Fluids*, 2010.
- [25] Szkuta, B.R., Sanabria, L.A., Dillon.T.S.: Electricity price short-term forecasting using artificial neural networks, *IEEE Trans. Power Sys.* 14, 851-857, 1999.
- [26] Vehviläinen, I., Pyykkönen, T.: Stochastic factor model for electricity spot price - the case of Nordic market, *Energy Economics*, forthcoming, 2005.
- [27] Voit J.: *The Statistical Mechanics of Financial Markets*. In: Springer-Verlag, Berlin, 2001.
- [28] Wang, A.J., Ramsay, B.: A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays, *Neurocomputing* 23, 47-57, 1998.

- [29] Weron, R., Misiorek, A.: Forecasting spot electricity prices with time series models. In: International Journal of Forecasting, Lodz, Poland, 2005.
- [30] Zivot E., Jiahui W.: Modeling Financial Time Series with S-Plus, 2002.
- [31] <http://en.wikipedia.org/wiki/Econophysics>
- [32] http://www.wordiq.com/definition/Efficient_market_hypothesis
- [33] <http://www.nordpool.com>
- [34] http://en.wikipedia.org/wiki/NASDAQ_OMX_Commodities_Europe
- [35] http://www.nsu.ru/matlab/Exponenta_RU/soft/matlab/stud1/main.asp.htm
(in Russian)