Jump Processes in Exchange Rates Modeling

Bunčák, Tomáš

Czech Technical University in Prague, Masaryk Institute of Advanced Studies

24 September 2013

Jump Processes in Exchange Rates Modeling

Working Paper

Tomáš Bunčák*

September 24, 2013

This text presents a study of various models based on jump processes in the context of foreign exchange (FX) rates modeling. Quality of FX rate log-returns fit is assessed for models such as *Merton* and *Kou jump-diffusions*, *normal inverse Gaussian*, *variance gamma*, and *Meixner*. The study is illustrated by simulation results that are provided for each of the models considered. Jump models are contrasted to the well-known (continuous) *Brownian motion* model.

Contents

*e-mail: buntom@gmail.com
Czech Technical University in Prague, Masaryk Institute of Advanced Studies

1 Introduction

In this text, we would like to provide an account of various *jump models* based on *Lévy processes* in the context of *foreign exchange (FX) rates* modeling. There are some publications considering jump models employed in FX modeling, among others [8, 15, 4, 18, 3, 17, 12]. What they basically conclude is that jumps are important part of FX rate models and that models based on random walks (e.g. Lévy processes) might be useful since it is rather difficult to model a systematic dependence on history of the process. Usually they come from a standard form of a model and modify it (e.g. include volatility clustering, heterogeneous jumps, etc.) to make the model more suitable for a specific FX rate modeling.

We give a comparative study of the widely used financial jump models in their *standard* setting. In the literature mentioned above, they work mostly with modifications of *jump-diffusion models*. Though we present these as well, we give examples of *infinite activity jump processes* too. It seems these are rarely used in FX rate modeling, although they provide good fit of FX rate returns, as we will see below. We compare these models regarding quality of FX rate (concretely we use EURUSD rate data) returns fit and we also present several results based on simulations.

The text is structured as follows. In Passage 2.1 we present FX rate modeling based on the standard *Brownian motion model* – just to show inferiority of a continuous model with respect to the jump models. In Passage 2.2.1 we give results corresponding to the *Merton jump-diffusion model*. Subsection 2.2.2 presents the *Kou jump-diffusion model*. Furthermore, we focus on infinite activity models; namely the *normal inverse Gaussian (NIG) model* (Passage 2.3.1), the *variance gamma (VG) model* (Passage 2.3.2), and finally the *Meixner model* (Passage 2.3.3). Passage 3 ends the treatise with some remarks and conclusions.

2 Comparison of Specific Lévy Processes in FX Modeling

We give examples of various Lévy processes that might be employed in foreign exchange rates modeling. We start with a simple *Brownian motion model*, then proceed to *jump-diffusion models*, and finally to models based on *infinite activity Lévy processes*. In all these examples we consider modeling of a logarithm exchange rate process since this is quite common in the financial literature. Logarithmic transformation makes the process more viable for modeling purposes. Hence by

$$r_t = \log(S_t),\ t \in \mathbf{T} = [0, T],\ T > 0,$$

we denote logarithmic FX rate process – considering $\{S_t\}$ as the original (non-transformed) FX rate process. Basically, we will try to model returns of this process (henceforth called *log-returns*) defined as $\Delta r_{t_i} = r_{t_i} - r_{t_i - \Delta t}$, where $t_i = i\Delta t$, $i = 1, \dots, \frac{T}{\Delta t}$, and $T \geq \Delta t > 0$ is such that $n := \frac{T}{\Delta t}$ is a natural number. Note that time step in which log-returns are sampled (from which parameters are estimated) is considered to be a unit time step, all the other Δt are supposed to be its multiples. Therefore there does not have to be Δt

in the estimation formulas below, although we write it in all the other formulas (where it is relevant).

2.1 Brownian Motion Model

In order to be able to see contrast with models involving jumps we provide modeling example of one of the simplest Lévy processes which is continuous, namely the *Brownian motion*.

Specification
In this case we have

$$r_t = \mu t + \sigma W_t,\ t \in \mathbf{T},$$

where by $\{W_t,\ t \in \mathbf{T}\}$ we denote *standard Brownian motion*, $\mu \in \mathbb{R}$ is a parameter called *drift*, $\sigma > 0$ is called *volatility*. By the *Itô formula* there is

$$\mathrm{d}S_t = \mathrm{d}e^{r_t} = S_t\big((\mu + \frac{1}{2}\sigma^2)\mathrm{d}t + \sigma \mathrm{d}W_t\big),\ t \in \mathbf{T},$$

thus the original FX rate follows the so-called *geometric Brownian motion*. It is clear that log-returns have the *normal (Gauss) distribution* $\mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$ with density of the form

$$f_{\Delta r_t}(x) = \phi(x; \mu\Delta t, \sigma^2\Delta t) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \exp\left(-\frac{(x-\mu\Delta t)^2}{2\sigma^2\Delta t}\right),\ x \in \mathbb{R}.$$

Estimation
Estimation of the parameters is quite simple in this setting – we employ the *method of moments* (MM). Estimates are given by

$$\hat{\mu} = \hat{\mathcal{M}}_r, \quad \hat{\sigma}^2 = \hat{\mathcal{V}}_r,$$

where $\hat{\mathcal{M}}_r$ denotes the *sample mean* and $\hat{\mathcal{V}}_r$ the *sample variance* of the sample of log-returns $\{\Delta r_{t_i},\ i = 1, \dots, n\}$.[1]

Statistics
In all of our modeling attempts we compare some sample statistics with their model counterparts. For this reason we state formulas for the theoretical (model) *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ for each of the models introduced in this text; in the Brownian setting simply

$$\mathcal{M}_r = \mu\Delta t, \quad \mathcal{V}_r = \sigma^2\Delta t, \quad \mathcal{S}_r = 0, \quad \mathcal{K}_r = 3.$$

[1] It is known that for the normal distribution, MM and *maximum-likelihood estimation* (MLE) yield technically the same estimates. Therefore, we do not needlessly employ MLE in this case.

Returns Modeling

In this passage, we provide results illustrating modeling capabilities of the *Gauss* (Brownian motion) model used for the FX rate log-returns[2]. We demonstrate the quality of fit of log-returns distribution and some simple "predictions" based on simulations. In what follows, analysis with the same structure can be found for any of the introduced models to make results comparable.

Firstly, let us show a comparison of empirical and model (probability) densities, namely Figure 1. In contrast with other (jump) models introduced within this text, we will see that the Gauss distribution does not fit very well. This might be seen also from Table 1, where we compare model and empirical values of some descriptive statistics for all the models considered herein. Moreover, we make use of some standard criteria for goodness-of-fit evaluation, namely the *root-mean-square error* (RMSE) and the *Bayesian information criterion* (BIC) – which is sometimes also called the *Schwarz criterion.* The former simply measures difference between the empirical and the model densities, namely

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{k} \left(f_{\Delta r_t}(x_i) - \hat{f}_{\Delta r_t}(x_i)\right)^2},$$

where $\{x_i,\, i = 1, \ldots, k\}$ is a given mesh of the returns density $f_{\Delta r_t}$ support and $\hat{f}_{\Delta r_t}$ denotes the empirical counterpart (EDF – empirical density function) of $f_{\Delta r_t}$. Clearly, smaller values of RMSE suggest a better fit. BIC is defined as

$$\mathrm{BIC} = 2L(\boldsymbol{\theta}) - \log(n)|\boldsymbol{\theta}|,$$

where $L(\boldsymbol{\theta})$ denotes the *log-likelihood function* (see (3)), n is the number of (log-returns) observations, and $|\boldsymbol{\theta}|$ stands for the number of parameters. Contrarily to RMSE, larger values of BIC should mean better fit quality. Overview of the values of criteria for different models might be found in Table 2. These values hint that Gauss distribution does not provide such a good fit of FX rate log-returns as the presented jump models.

Model	$\mathcal{M}^r$	$\mathcal{V}^r$	$\mathcal{S}^r$	$\mathcal{K}^r$
Gauss	-9.5991e-007	1.8518e-006	0	3
Merton JD	-5.2788e-006	1.7524e-006	-0.02194	7.07
Kou JD	-9.5971e-007	1.7589e-006	-0.078549	10.2913
NIG	-9.6014e-007	1.8683e-006	-0.09745	12.7511
VG	-2.6699e-006	1.6683e-006	0.042273	6.8322
Meixner	-9.6475e-007	1.8256e-006	-0.07506	10.3032
Empirical	$\hat{\mathcal{M}}^r$	$\hat{\mathcal{V}}^r$	$\hat{\mathcal{S}}^r$	$\hat{\mathcal{K}}^r$
	-9.5991e-007	1.8518e-006	0.094547	13.6506

Table 1: Comparison of empirical and model statistics of log-returns – mean, variance, skewness, kurtosis; fitted on EURUSD [1h returns 2005 – 2012]

[2]We work with EURUSD exchange rate data sampled hourly for the period 01/01/2005 – 27/10/2012.

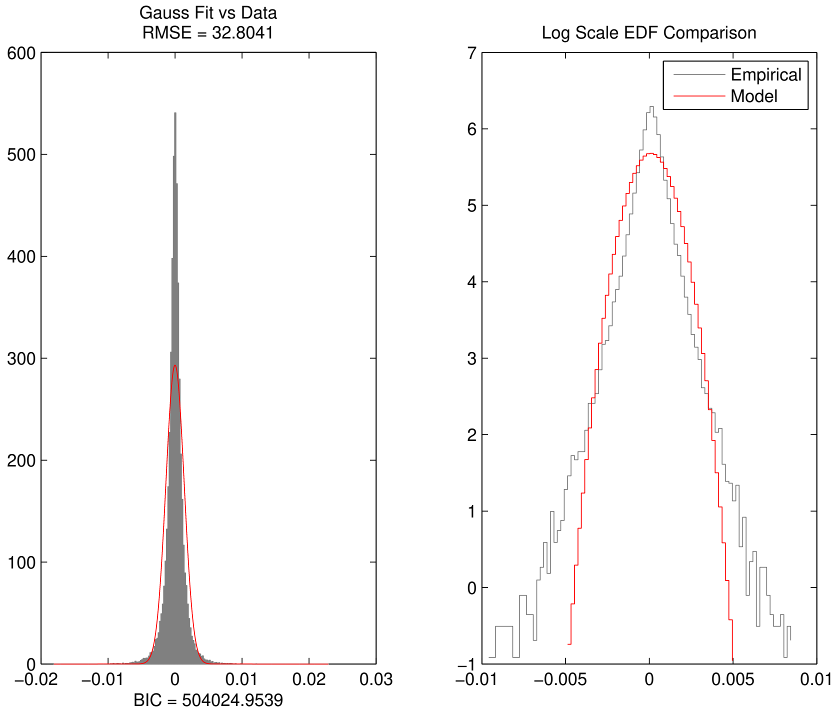

Figure 1: Comparison of empirical and model densities – Gauss model; EURUSD [1h returns 2005 – 2012]

*Model**Criterion*	RMSE	BIC
Gauss	32.8041	504024.9539
Merton JD	6.4403	519362.9572
Kou JD	3.8792	519856.9482
NIG	2.3274	520121.4649
VG	7.4299	519486.5831
Meixner	1.6573	520129.278

Table 2: Goodness-of-fit criteria (RMSE and BIC) for different models; fitted on EURUSD [1h returns 2005 – 2012]

Furthermore, we present a point prediction gained by the Gauss model simulation[3] of log-returns. For the distribution of the simulated FX rate, see Figure 2. Further on we will see that these “predictions” are quite similar for all of the models in this text. Though this is not surprising, considering the *central limit theorem* and the fact that during the simulations, we make sums if independent identically distributed random variables with similar means and variances.

Finally, let us present more results based on simulations. To get a rough image of model trajectories, see Figure 3, where we display 5 simulated trajectories opposed to the

[3]Each simulation is conducted by sampling of 10^6 values/trajectories.

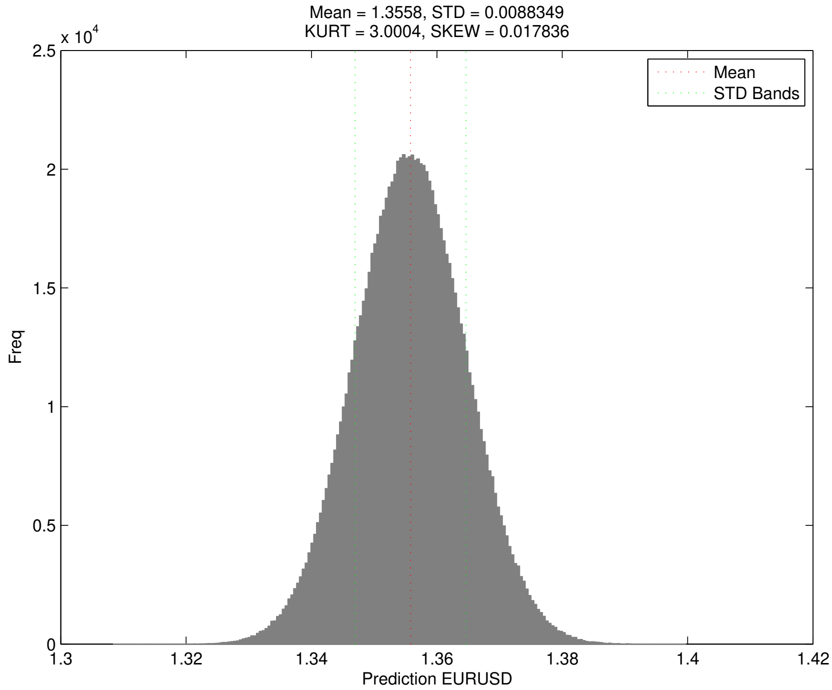

Figure 2: Simulated point prediction, 24h period ahead – Gauss model; fitted on EURUSD [1h returns 2005 – 2012]

real FX rate trajectory to whose returns model has been calibrated. Getting back to a predictive kind of analysis, we try to fit model on a shorter period of time (in fact half of the sample) and then simulate the FX rate process for 80 (trading hour) periods ahead and compare this with the observed trajectory. In Figure 4, we show the real FX rate trajectory together with the mean of simulated trajectories and standard deviation bands around this mean. Here, the displayed RMSE is calculated from differences between simulated and real FX rate values. Note that again, as in the point prediction case, results of this “prediction” will be quite comparable throughout the different models with almost no regard to the quality of fit.

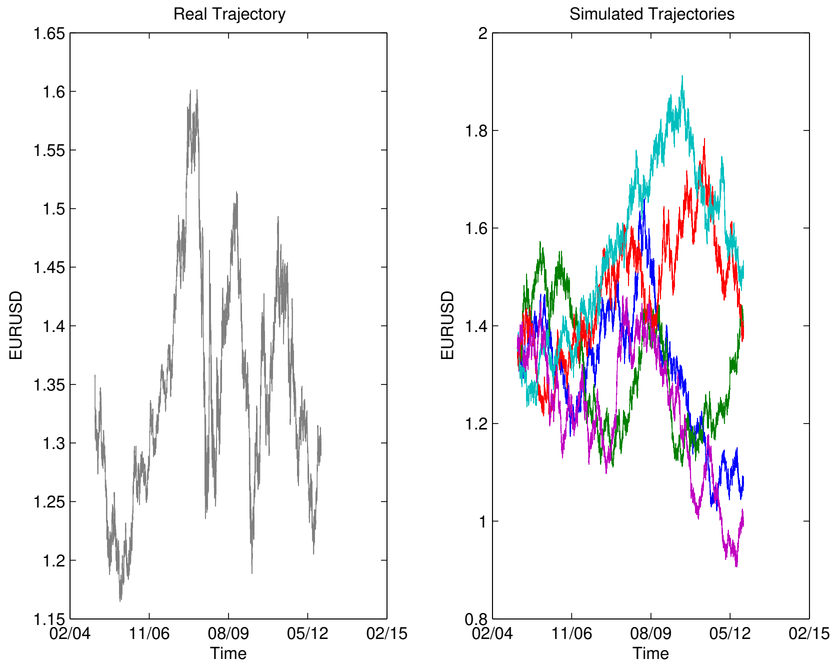

Figure 3: Real vs simulated trajectories – Gauss model; EURUSD

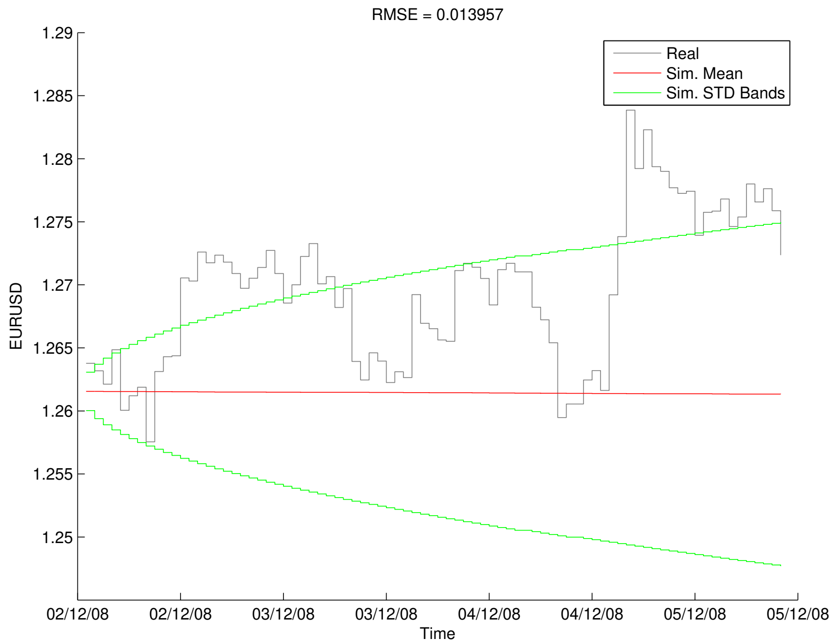

Figure 4: Out-of-sample simulation vs real trajectory – Gauss model; EURUSD

2.2 Jump-Diffusion Models

Jump-diffusion (JD) models form one of the subgroups of Lévy process based models. Within this text, one may consider it as a direct generalization of Brownian motion models since it is a superposition of a (drifted) Brownian motion process and a *compound Poisson process*. In these type of models jumps are "rare" events with a prescribed distribution governing their occurrence and magnitude. See [5, Subsection 4.1] for a discussion of models based on Lévy processes.

We introduce two models of this type, namely the *Merton JD model* (see Passage 2.2.1) and the *Kou JD model* (see Passage 2.2.2). They differ only in distribution of jumps as we will see below.

2.2.1 Merton JD Model

Merton JD is a JD model with jump sizes that are normally distributed. This model was introduced in [16].

Specification
Logarithmic FX rate process follows

$$r_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i,\ t \in \mathbf{T}, \tag{1}$$

where $\mu \in \mathbb{R}$ is *drift*, $\sigma > 0$ *volatility*, $\{N_t,\, t \in \mathbf{T}\}$ is a *Poisson process* with *(jump) intensity* $\lambda > 0$, and $\{Y_i,\, i = 1, 2, \dots\}$ are independent identically distributed (iid) random variables with normal distribution $\mathcal{N}(\xi, \tau^2)$ – hence $\xi \in \mathbb{R}$ is the *mean of jumps* parameter and $\tau^2 > 0$ is the *variance of jumps* parameter. Remind that by $\{W_t\}$ we denote standard Brownian motion. Note that $\{W_t\}$, $\{N_t\}$ and $\{Y_i\}$ are all mutually independent. By the *Itô formula for jump-diffusion processes*, see [19, Theorem 1.14] for instance, we have

$$\begin{aligned} \mathrm{d}S_t = \mathrm{d}e^{r_t} &= S_t\left((\mu + \tfrac{1}{2}\sigma^2)\mathrm{d}t + \sigma \mathrm{d}W_t\right) + \\ &+ S_{t-}\int_{\mathbb{R}} (e^z - 1) N(\mathrm{d}t, \mathrm{d}z),\ t \in \mathbf{T}, \end{aligned}$$

where $N(\mathrm{d}t, \mathrm{d}z)$ is the *Poisson random measure* associated with the Poisson process $\{N_t\}$, see [5, Subsection 2.6] for instance. *Lévy measure* of this process is given by

$$\nu(\mathrm{d}z) = \phi(z; \xi, \tau^2)\mathrm{d}z = \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(z-\xi)^2}{2\tau^2}\right) \mathrm{d}z \text{ on } \mathbb{R}.$$

Density of the log-returns is given by

$$\begin{cases} f_{\Delta r_t}(x) = e^{-\lambda \Delta t} \displaystyle\sum_{k=0}^{\infty} \left\{ \frac{(\lambda \Delta t)^k}{k!} \exp\left(-\frac{(x - \mu\Delta t - k\xi)^2}{2(\sigma^2 \Delta t + k\tau^2)}\right) \cdot \right. \\ \qquad \left. \cdot \ (2\pi(\sigma^2\Delta t + k\tau^2))^{-\frac{1}{2}} \right\},\ x \in \mathbb{R}. \end{cases} \tag{2}$$

Estimation

To estimate Merton JD parameters we employ *maximum-likelihood estimation* (MLE) technique. We want to estimate vector of parameters $\boldsymbol{\theta} = (\mu, \sigma, \lambda, \xi, \tau)$ by maximization of the *log-likelihood function*, namely

$$\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\theta}) = \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \sum_{i=1}^{n} \log(f_{\Delta r_t}(\Delta r_{t_i}; \boldsymbol{\theta})), \tag{3}$$

where $f_{\Delta r_t}(x; \boldsymbol{\theta})$ is the density given by (2) – corresponding to specific vector of parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. By $\boldsymbol{\Theta}$ we denote the *parameter space* defined with regards to conditions on parameters introduced with the model above.[4] As we can see in (2), expression for the density involves infinite summation. However, this is not a big issue since the convergence is sufficiently fast, so we can approximate the infinite sum by the sum of first $k_{\mathrm{MAX}} = 100$ elements for instance. The last thing is to find an appropriate initial estimate for the maximizing procedure. This is ensured by implementation of the *expectation-maximization* algorithm described in [6].

Statistics

Statistics of the Merton JD model – namely *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ – are as follows

$$\mathcal{M}_r = (\mu + \lambda\xi)\Delta t, \qquad \mathcal{V}_r = (\sigma^2 + \lambda\tau^2 + \lambda\xi^2)\Delta t,$$
$$\mathcal{S}_r = \frac{(3\tau^2\xi + \xi^3)\lambda}{(\sigma^2+\lambda\tau^2+\lambda\xi^2)^{\frac{3}{2}}\sqrt{\Delta t}}, \quad \mathcal{K}_r = \frac{(3\tau^4+6\xi^2\tau^2+\xi^4)\lambda}{(\sigma^2+\lambda\tau^2+\lambda\xi^2)^2\Delta t} + 3.$$

Returns Modeling

Following the structure of the corresponding passage for the previously introduced Brownian motion (Gauss) model, we present analogous group of results. Quality of (log-returns distribution) fit might be assessed from Figure 5.

We see that the fit is better than in the Gaussian case, but some of the other jump models surpass even this fit – as we will find further on.

Onwards, let us illustrate model trajectories – see Figure 6. We also provide sketch of a distribution of simulated “point prediction” of the FX rate value 24 hour period ahead, see Figure 7. Finally, we show Figure 8 of the out-of-sample (80 hourly periods of simulation, fitted on the half of the sample) simulated trajectories’ mean and corresponding standard deviation bands. As we have said before, these will not differ much throughout the different models considered herein.

[4] In each of the models we need to consider specific parameter space constraints, of course.

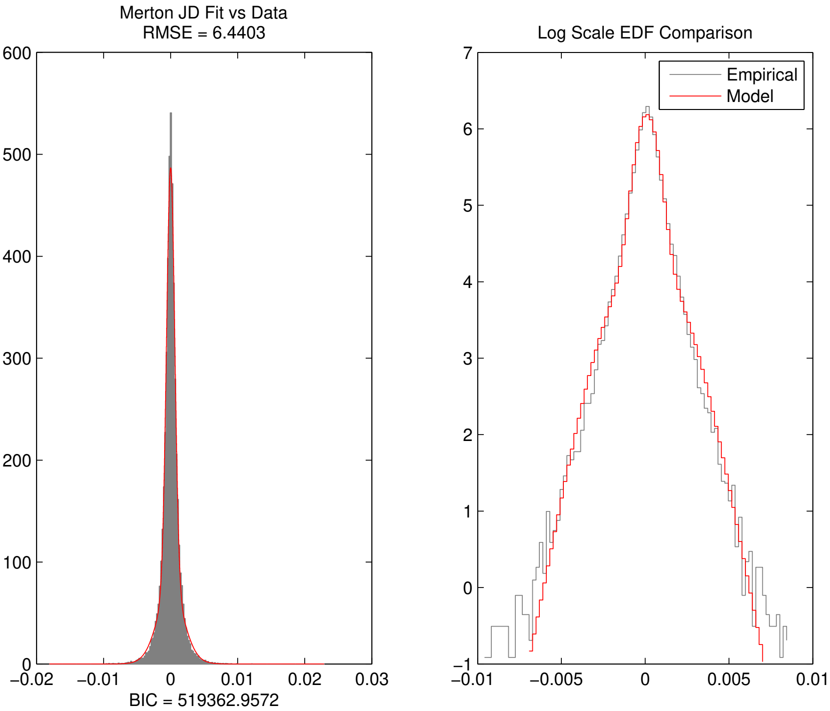

Figure 5: Comparison of empirical and model densities – Merton JD model; EURUSD [1h returns 2005 – 2012]

2.2.2 Kou JD Model

Here we have an example of another jump-diffusion model with a different distribution of jumps sizes – a mixture of exponential distributions. The model was introduced in [9].

Specification

Logarithmic FX rate follows the same type of the process as in (1). The only difference is that jump sizes $\{Y_i\}$ have *double exponential distribution* – that is a mixture of two exponential distributions, one for negative and one for positive jumps. So density of the jump size is given as

$$f_Y(y) = p\lambda_+ e^{-\lambda_+ y}\mathbb{1}_{[y\geq 0]} + (1-p)\lambda_- e^{-\lambda_- |y|}\mathbb{1}_{[y<0]},\ y \in \mathbb{R},$$

where $p \in [0,1]$, $\lambda_+ > 0$ is the reciprocal of positive jumps mean and $\lambda_- > 0$ is the reciprocal of absolute value of negative jumps mean. From this follows the *Lévy measure*, namely

$$\nu(\mathrm{d}z) = f_Y(z)\mathrm{d}z = \big(p\lambda_+ e^{-\lambda_+ z}\mathbb{1}_{[z\geq 0]} + (1-p)\lambda_- e^{-\lambda_- |z|}\mathbb{1}_{[z<0]}\big)\mathrm{d}z \text{ on } \mathbb{R}.$$

Density of the log-returns is not available in a closed form, however, we may express it in an approximate shape. This will be sufficient for our modeling purposes. For this reason

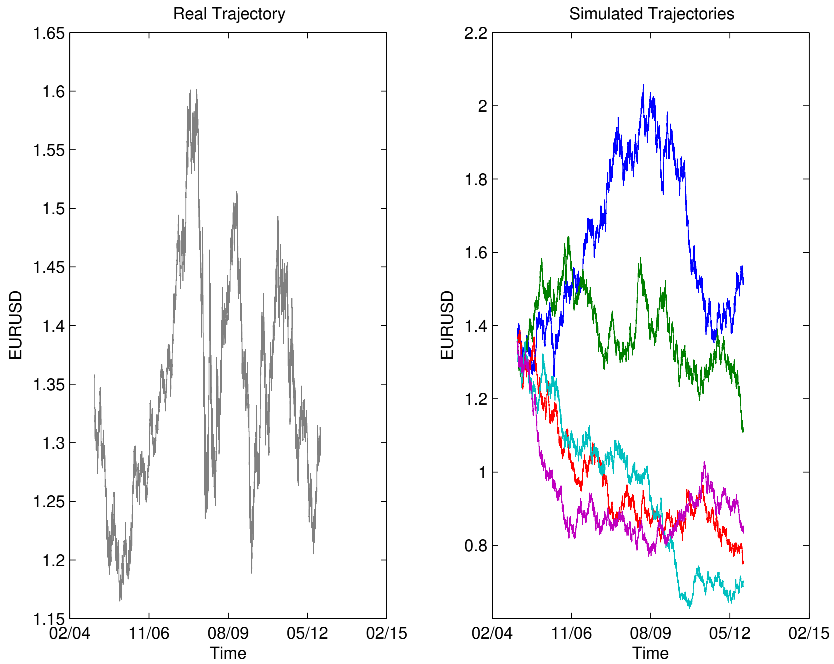

Figure 6: Real vs simulated trajectories – Merton JD model; EURUSD

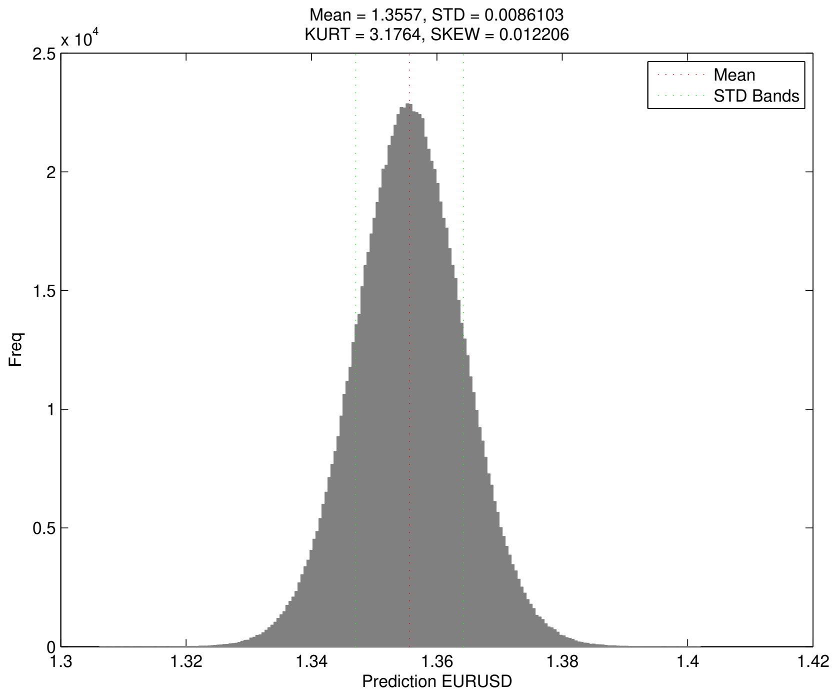

Figure 7: Simulated point prediction, 24h period ahead – Merton JD model; fitted on EURUSD [1h returns 2005 – 2012]

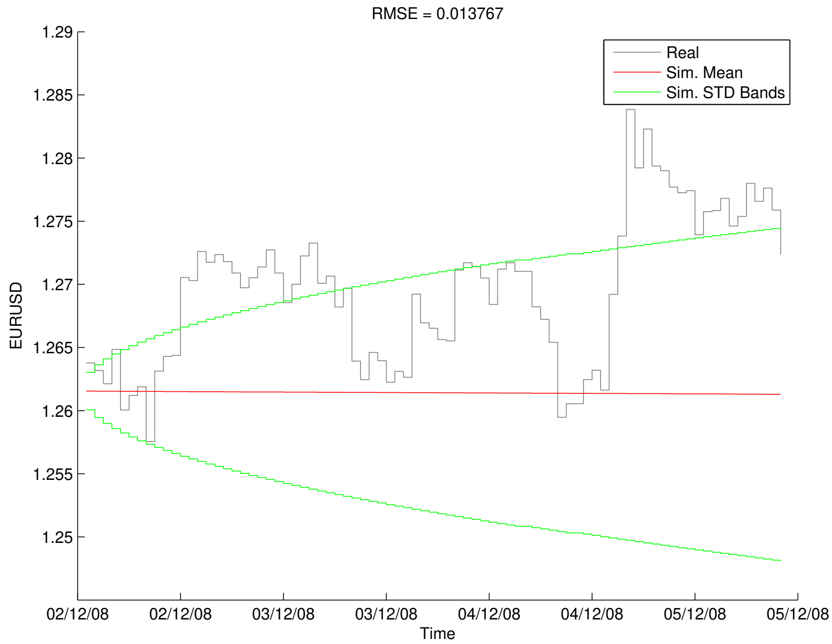

Figure 8: Out-of-sample simulation vs real trajectory – Merton JD model; EURUSD

we neglect probability of more than one jumps in a single period log-return. Then by the formulas for a mixture of normal and exponential distribution (see [9] for instance) we get

$$
(4) \qquad \begin{cases} f_{\Delta r_t}(x) & \doteq & \lambda\Delta t\Big\{ p\lambda_+ e^{\frac{\sigma^2\lambda_+^2\Delta t}{2} - (x-\mu\Delta t)\lambda_+} \cdot \Phi\big(\frac{x-\mu\Delta t-\sigma^2\lambda_+\Delta t}{\sigma\sqrt{\Delta t}}\big) + \\ & + & q\lambda_- e^{\frac{\sigma^2\lambda_-^2\Delta t}{2} + (x-\mu\Delta t)\lambda_-} \cdot \Phi\big(-\frac{x-\mu\Delta t+\sigma^2\lambda_-\Delta t}{\sigma\sqrt{\Delta t}}\big)\Big\} + \\ & + & \frac{1-\lambda\Delta t}{\sigma\sqrt{\Delta t}}\phi\big(\frac{x-\mu\Delta t}{\sigma\sqrt{\Delta t}}\big), \; x \in \mathbb{R}, \end{cases}
$$

where $q = 1 - p$, Φ and ϕ are the cumulative distribution function and the probability density function of the standard normal distribution, respectively.

Estimation

Estimation is performed by MLE as in the Merton JD model. Here we want to estimate vector of parameters $\boldsymbol{\theta} = (\mu, \sigma, \lambda, p, \lambda_+, \lambda_-)$. Naturally, we plug (4) in (3) and run the maximization procedure. To determine initial estimates for the MLE maximization (3) we use "intuitive guesses". More specifically, jumps are considered to be returns with absolute value above some high quantile of the absolute log-returns. Initial p is then estimated as the ratio of positive jumps count to the number of all the jumps observed. Intensity of jumps is guessed from data and may be adjusted should the maximization procedure has some problems with convergence. Parameters λ_+ and λ_- are given as reciprocals of sample means of positive and (absolute value of) negative jumps, respectively. Finally, drift μ and volatility σ are initiated by sample mean and sample standard deviation of non-jump observations of log-returns.

Statistics

Statistics of the Kou JD model – namely *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ – are as follows

$$\mathcal{M}_r = (\mu + p\frac{\lambda}{\lambda_+} - (1-p)\frac{\lambda}{\lambda_-})\Delta t, \quad \mathcal{V}_r = (\sigma^2 + 2p\frac{\lambda}{\lambda_+^2} + 2(1-p)\frac{\lambda}{\lambda_-^2})\Delta t,$$

$$\mathcal{S}_r = \frac{6(p/\lambda_+^3 - (1-p)/\lambda_-^3)\lambda\Delta t}{\mathcal{V}_r^{\frac{3}{2}}}, \qquad \mathcal{K}_r = \frac{24(p/\lambda_+^4 + (1-p)/\lambda_-^4)\lambda\Delta t}{\mathcal{V}_r^2} + 3.$$

Returns Modeling

Quality of the Kou JD model log-returns fit is displayed in Figure 9. We see that for these data, it performs slightly better than the Merton JD model.

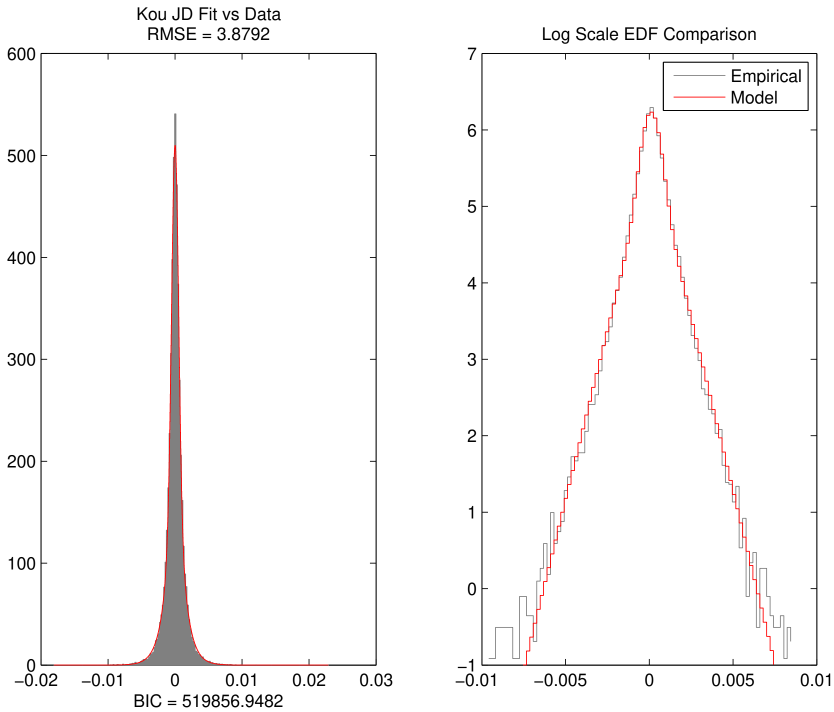

Figure 9: Comparison of empirical and model densities – Kou JD model; EURUSD [1h returns 2005 – 2012]

Again, we present results based on simulations. Namely an illustration of the model trajectories in Figure 10, simulated point "prediction" of the FX rate in Figure 11, and mean and standard deviation bands of simulated trajectories in Figure 12. Any comments follow the line of those presented for the previously introduced models.

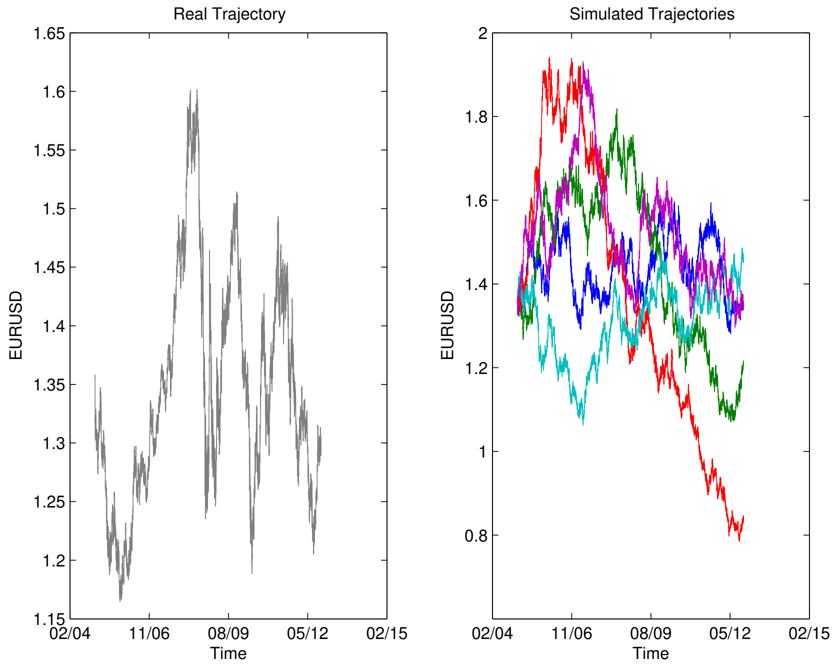

Figure 10: Real vs simulated trajectories – Kou JD model; EURUSD

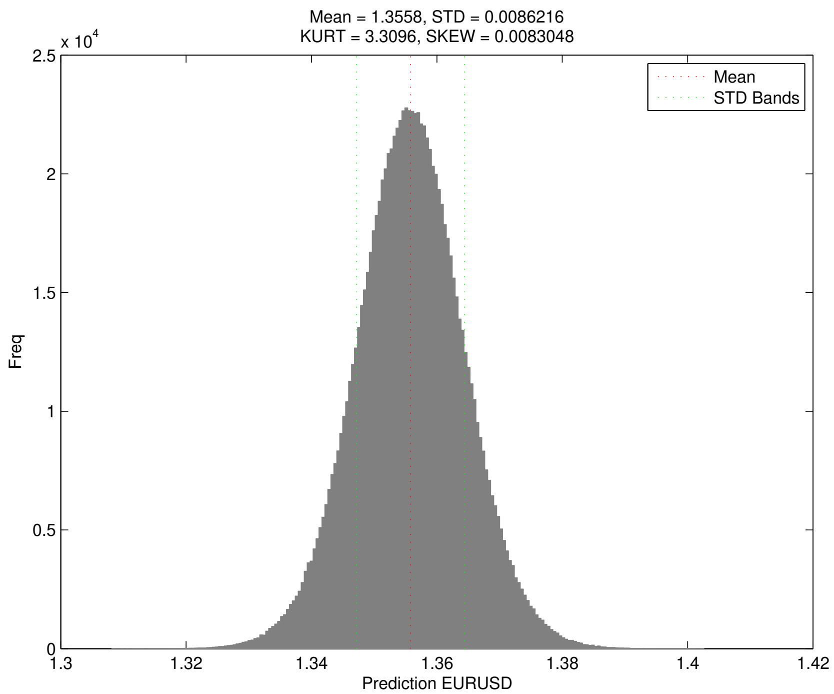

Figure 11: Simulated point prediction, 24h period ahead – Kou JD model; fitted on EURUSD [1h returns 2005 – 2012]

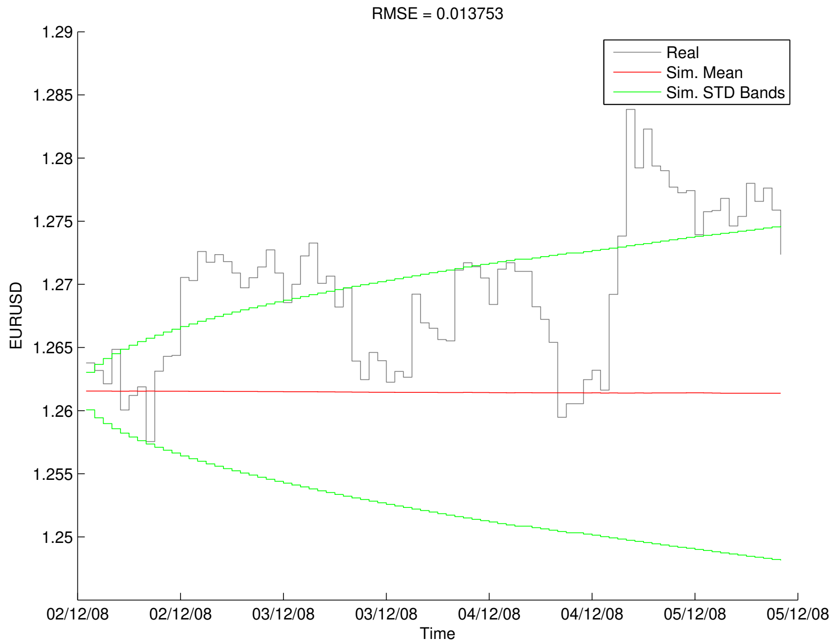

Figure 12: Out-of-sample simulation vs real trajectory – Kou JD model; EURUSD

2.3 Infinite Activity Models

Infinite activity (Lévy) models is a subclass of models based on Lévy processes which include infinite number of jumps in each (finite) interval. Hence it is not necessary to introduce Brownian motion component in these models since even a small interval behavior is well described by the infinite activity jump component. However, some of these models may be expressed by the so-called *Brownian subordination* – i.e. as a time-changed Brownian motion. This we will see in the examples of *normal inverse Gaussian* (see Passage 2.3.1) and *variance gamma* (see Passage 2.3.2) models. Furthermore, we demonstrate an application of another model used in finance, namely the *Meixner model* (see Passage 2.3.3).

2.3.1 Normal Inverse Gaussian Model

Normal inverse Gaussian (NIG) model is a member of the class of *generalized hyperbolic models*, i.e. models based on processes yielded by Brownian subordination with the *generalized inverse Gaussian subordinator*. For more information about this class see [5, Section 4.6] for instance. We will focus on the NIG model which was introduced to finance in [2].

Specification

Logarithmic FX rate follows the process

$$r_t = \mu t + \beta Z_t + W_{Z_t},\ t \in \mathbf{T}. \tag{5}$$

Parameters of the NIG model are $\mu \in \mathbb{R}$, $\alpha > 0$, $\delta > 0$, and $0 < |\beta| < \alpha$. Process $\{W_t\}$ is a standard Brownian motion. By $\{Z_t\}$ we denote the *inverse Gaussian subordinator*, that is a (non-decreasing non-negative) process whose increments are governed be the *inverse Gaussian law*. This means that $\{\Delta Z_{t_i} = Z_{t_i} - Z_{t_i - \Delta t},\ i = 1, \dots, n\}$ has a distribution $IG(\frac{\delta}{\gamma}\Delta t, \delta^2(\Delta t)^2)$, where $\gamma = \sqrt{\alpha^2 - \beta^2}$ and $IG(\xi, \lambda)$ for any $\xi > 0$ and $\lambda > 0$ has a probability density of the form

$$f_{\mathrm{IG}}(z; \xi, \lambda) = \sqrt{\frac{\lambda}{2\pi z^3}} \exp\left(-\frac{\lambda(z-\xi)^2}{2\xi^2 z}\right),\ z > 0,$$

and, of course, $f_{\mathrm{IG}} \equiv 0$ for $z \leq 0$. Interpretation of the NIG parameters is following. Parameter α adjusts tail behavior, that is steepness of the NIG returns distribution – larger α is, lighter are the tails of the distribution. Parameter β controls the skewness of the returns distribution; $\beta < 0$ produces left-skewed densities, $\beta > 0$ right-skewed. Standardly, μ is the location (or drift) parameter. Finally, δ plays a similar role to σ in the previously introduced models, that is representing measure of the volatility of the returns; higher δ leads to more volatile returns.

Furthermore, we introduce formula for the density of log-returns

$$\left\{\begin{array}{rcl} f_{\Delta r_t}(x) & = & \frac{\alpha\delta\Delta t}{\pi} \exp(\lambda\gamma\Delta t + \beta(x - \mu\Delta t)) \cdot \\ & \cdot & \frac{K_1(\alpha\sqrt{\delta^2(\Delta t)^2 + (x-\mu\Delta t)^2})}{\sqrt{\delta^2(\Delta t)^2 + (x-\mu\Delta t)^2}},\ x \in \mathbb{R}, \end{array}\right. \tag{6}$$

where K_u denotes the *modified Bessel function of the second kind and index* u, see [1] for instance.

We also give the *Lévy measure* for this model, namely

$$\nu(\mathrm{d}z) = \frac{\alpha\delta}{\pi} \frac{\exp(\beta z) K_1(\alpha|z|)}{|z|} \mathrm{d}z \text{ on } \mathbb{R} \setminus \{0\}.$$

Estimation

Estimation of parameters is performed by MLE (3), considering that we have a closed formula for the density of log-returns (6). Here we want to find an optimal vector of parameters $\boldsymbol{\theta} = (\alpha, \beta, \mu, \sigma)$. Again, we need to start MLE maximization procedure with some initial values of parameters. This might be done with a MM estimate given by the following formulas

$$\hat{\alpha} = 3\hat{\rho}^{\frac{1}{2}}(\hat{\rho} - 1)^{-1}\hat{\mathcal{V}}_r^{-\frac{1}{2}}|\hat{\mathcal{S}}_r|^{-1}, \quad \hat{\beta} = 3(\hat{\rho} - 1)^{-1}\hat{\mathcal{V}}_r^{-\frac{1}{2}}\hat{\mathcal{S}}_r^{-1},$$
$$\hat{\mu} = \hat{\mathcal{M}}_r - 3\hat{\rho}^{-1}\hat{\mathcal{V}}_r^{\frac{1}{2}}\hat{\mathcal{S}}_r^{-1}, \qquad \hat{\delta} = 3\hat{\rho}^{-1}(\hat{\rho} - 1)^{\frac{1}{2}}\hat{\mathcal{V}}_r^{\frac{1}{2}}|\hat{\mathcal{S}}_r|^{-1},$$

where $\hat{\rho} = 3(\hat{\mathcal{K}}_r - 3)\hat{\mathcal{S}}_r^{-2} - 4$, and $\hat{\mathcal{M}}_r$, $\hat{\mathcal{V}}_r$, $\hat{\mathcal{S}}_r$, $\hat{\mathcal{K}}_r$ are *sample mean*, *sample variance*, *sample skewness*, and *sample kurtosis* (of the log-returns), respectively. Let us note that for this MM estimation there has to be $3(\hat{\mathcal{K}}_r - 3) > 5\hat{\mathcal{S}}_r^2 > 0$. These formulas are taken from [7].

Statistics

Statistics of the NIG model – namely *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ – are as follows

$$\mathcal{M}_r = (\mu + \tfrac{\bar{\kappa}\delta}{\sqrt{1-\bar{\kappa}^2}})\Delta t, \qquad \mathcal{V}_r = \delta^2/(\bar{\alpha}(1-\bar{\kappa}^2)^{\frac{3}{2}})\Delta t,$$
$$\mathcal{S}_r = 3\bar{\kappa}/(\sqrt{\bar{\alpha}\Delta t}(1-\bar{\kappa}^2)^{\frac{1}{4}}), \quad \mathcal{K}_r = 3(4\bar{\kappa}^2+1)/(\bar{\alpha}\Delta t(1-\bar{\kappa}^2)^{\frac{1}{2}}) + 3,$$

where $\bar{\alpha} = \alpha\delta$, $\bar{\kappa} = \beta/\alpha$.

Returns Modeling

NIG model fits the log-returns as displayed in Figure 13. This seems to be better than both jump-diffusion models presented above.

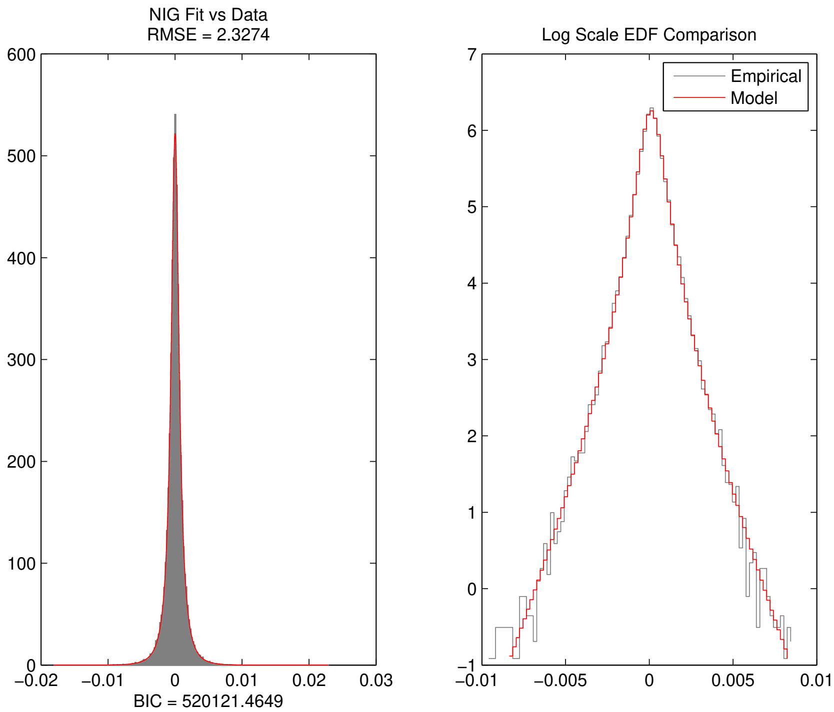

Figure 13: Comparison of empirical and model densities – NIG model; EURUSD [1h returns 2005 – 2012]

Standardly, let us present the simulation results. Sample trajectories are depicted in Figure 14. Point "prediction" can be seen in Figure 15. Finally, mean of simulated trajectories with standard deviation bands can be found in Figure 16.

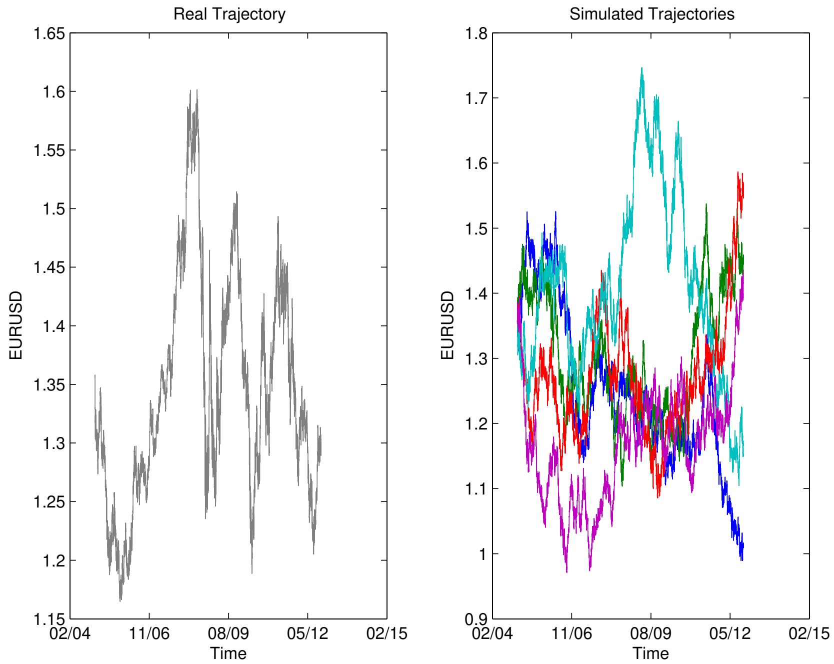

Figure 14: Real vs simulated trajectories – NIG model; EURUSD

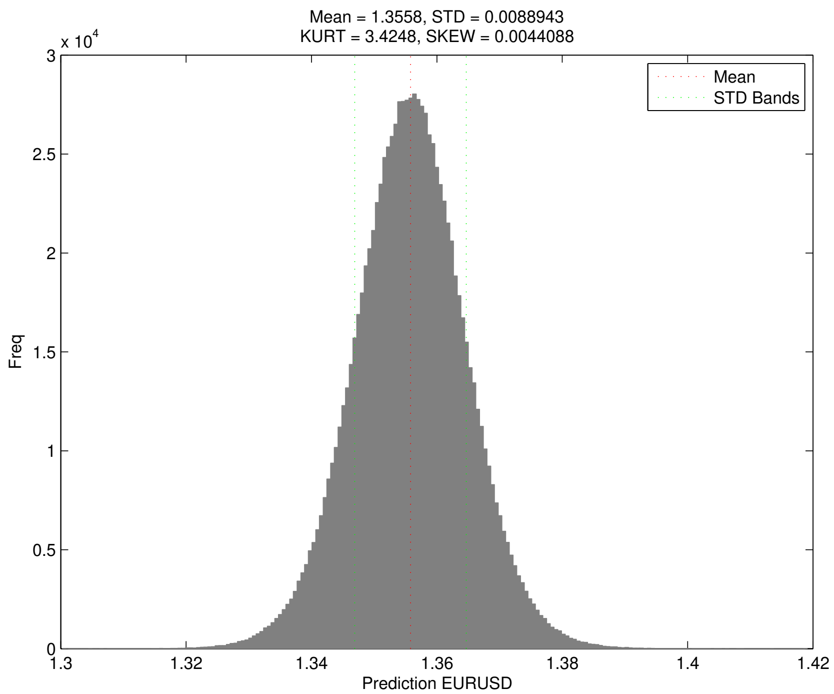

Figure 15: Simulated point prediction, 24h period ahead – NIG model; fitted on EURUSD [1h returns 2005 – 2012]

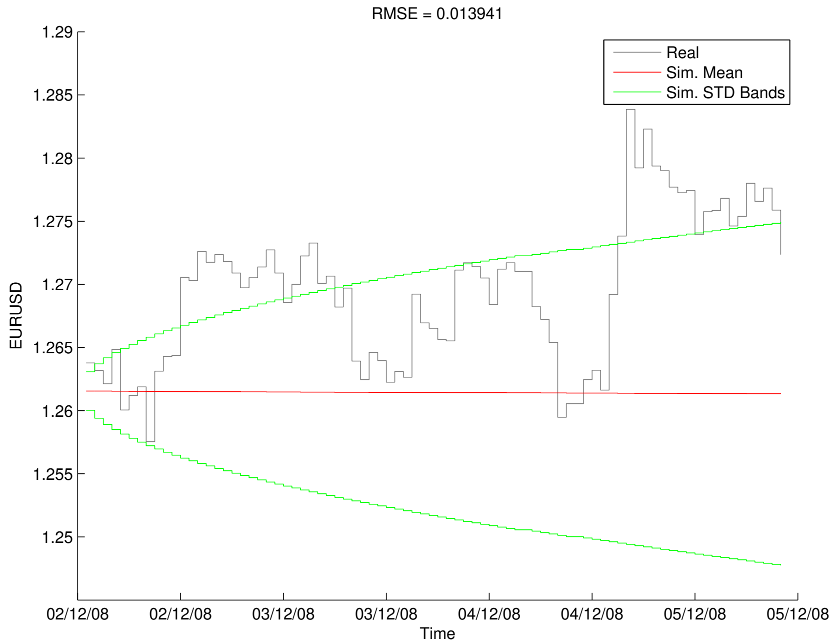

Figure 16: Out-of-sample simulation vs real trajectory – NIG model; EURUSD

2.3.2 Variance Gamma Model

Variance gamma (VG) process is an example of a *generalized tempered stable process*. A subclass of these processes that groups processes representable by Brownian motion subordination is called CGMY and VG process belongs to this subclass. Further information about these processes might be found in [5, Section 4.5] for instance. Introduction of the VG model to finance dates back to [10].

Specification

The model process for the FX rate is similar to the one considered in the NIG model (see (5)). In the VG model there holds

$$r_t = \mu t + \beta Z_t + \sigma W_{Z_t},\ t \in \mathbf{T},$$

where $\mu \in \mathbb{R}$, $\beta \in \mathbb{R}$, and $\sigma > 0$. Subordinating process $\{Z_t\}$ is given by a *gamma process*, i.e. process whose increments $\{\Delta Z_{t_i} = Z_{t_i} - Z_{t_i - \Delta t},\ i = 1, \ldots, n\}$ follow the *gamma distribution* $\Gamma(\Delta t/\kappa, \kappa)$. Parameter $\kappa > 0$ defines variance of the subordinator. Standardly, μ dictates drift of the process in a calendar time. We may think of β as a drift of the process in a *business time*, i.e. time given by the subordinator. Finally, σ controls volatility of the process.

Density of $\Gamma(k, \theta)$ for any $k > 0$ and $\theta > 0$ is given by

$$f_\Gamma(z; k, \theta) = \frac{1}{\theta^k \Gamma(k)} z^{k-1} \exp\left(-\frac{z}{\theta}\right),\ z > 0,$$

naturally $f_\Gamma \equiv 0$ for $z \le 0$; by Γ in the formula we denote the well-known *gamma function*. Density of the log-returns is then as follows

$$\text{(7)} \qquad \begin{cases} f_{\Delta r_t}(x) &= \frac{2(\beta^2+2\sigma^2/\kappa)^{\frac{1}{4}-\frac{\Delta t}{2\kappa}}}{\kappa^{\frac{\Delta t}{\kappa}}\sqrt{2\pi}\sigma\Gamma(\Delta t/\kappa)}|x-\mu\Delta t|^{\frac{\Delta t}{\kappa}-\frac{1}{2}} \cdot e^{\beta/\sigma^2(x-\mu\Delta t)}. \\ &\cdot \quad K_{\frac{\Delta t}{\kappa}-\frac{1}{2}}\left(\frac{\sqrt{\beta^2+2\sigma^2/\kappa}}{\sigma^2}|x-\mu\Delta t|\right), \; x \in \mathbb{R}, \end{cases}$$

where K_u denotes the *modified Bessel function of the second kind and index* u, see [1] if needed.

Moreover, *Lévy measure* corresponding to the VG model is expressed by

$$\nu(\mathrm{d}z) = \frac{1}{\kappa|z|}\exp\left(\frac{\beta}{\sigma^2}z - \frac{\sqrt{\beta^2+2\sigma^2/\kappa}}{\sigma^2}|z|\right)\mathrm{d}z \text{ on } \mathbb{R}\setminus\{0\}.$$

Estimation

To estimate VG model parameters, we employ MLE again. So we want to estimate $\boldsymbol{\theta} = (\mu, \beta, \sigma, \kappa)$ by (3) where we use (7). We initiate the MLE maximization procedure with a MM estimate as in the NIG model. Here, the MM estimate is given by the following procedure:

1. Find (numerically) a solution ε^* to
$$\frac{\varepsilon(3+2\varepsilon)^2}{(1+4\varepsilon+2\varepsilon^2)(1+\varepsilon)} = \frac{3\hat{\mathcal{S}}_r^2}{\hat{\mathcal{K}}_r - 3}.$$
2. Compute MM estimates by
$$\hat\sigma^2 = \hat{\mathcal{V}}_r/(1+\varepsilon^*), \quad \hat\kappa = \tfrac{1}{3}(\hat{\mathcal{K}}_r - 3)\tfrac{(1+\varepsilon^*)^2}{1+4\varepsilon^*+2\varepsilon^{*2}},$$
$$\hat\beta = \tfrac{\hat{cm}_{r,3}}{\hat\sigma^2\hat\kappa}\tfrac{1}{3+2\varepsilon^*}, \qquad \hat\mu = r_T/n - \hat\beta,$$
where $\hat{cm}_{r,3}$ denotes the *third sample central moment* of the log-returns.

Recall that by $\hat{\mathcal{V}}_r$ we mean *sample variance*, by $\hat{\mathcal{S}}_r$ *sample skewness*, and by $\hat{\mathcal{K}}_r$ *sample kurtosis* (of the log-returns). This MM estimate is adopted from [14].

Statistics

Statistics of the VG model – namely *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ – are as follows

$$\mathcal{M}_r = (\mu+\beta)\Delta t, \qquad \mathcal{V}_r = (\sigma^2+\beta^2\kappa)\Delta t,$$
$$\mathcal{S}_r = \tfrac{(3\sigma^2\beta\kappa+2\beta^3\kappa^2)\Delta t}{\mathcal{V}_r^{\frac{3}{2}}}, \quad \mathcal{K}_r = \tfrac{(3\sigma^4\kappa+6\beta^4\kappa^3+12\sigma^2\beta^2\kappa^2)\Delta t}{\mathcal{V}_r^2} + 3.$$

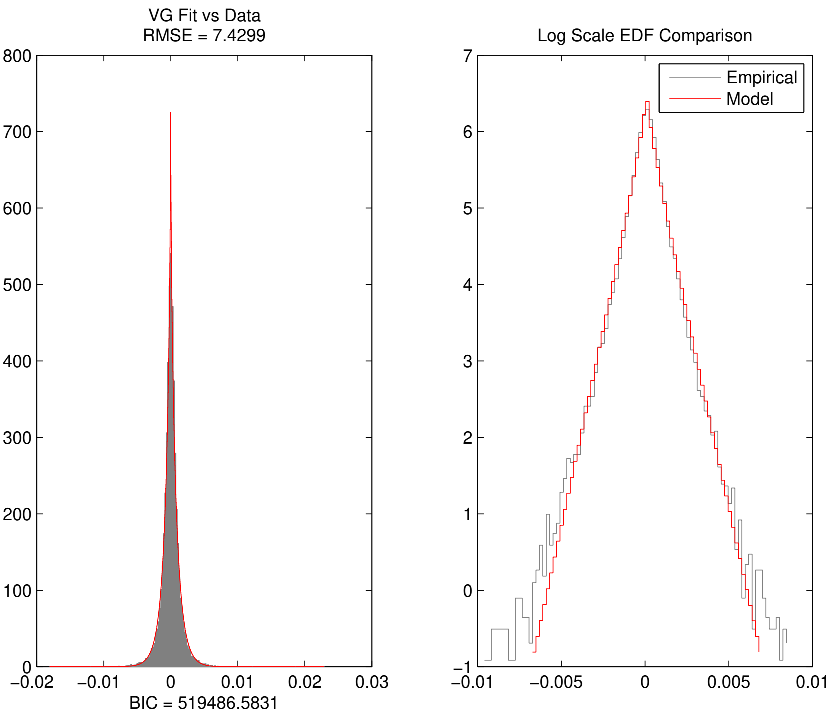

Figure 17: Comparison of empirical and model densities – VG model; EURUSD [1h returns 2005 – 2012]

Returns Modeling

VG model provides a fit which is illustrated in Figure 17. We see that the quality is worse than quality of fits provided by all of the models above, except for the Gauss model (which is worse than all the jump models).

Furthermore, we present simulation results again. Sample trajectories of the VG model are depicted in Figure 18. FX rate simulated "prediction" distribution can be seen in Figure 19. Lastly, mean and standard deviation bands of simulated trajectories might be found in Figure 20.

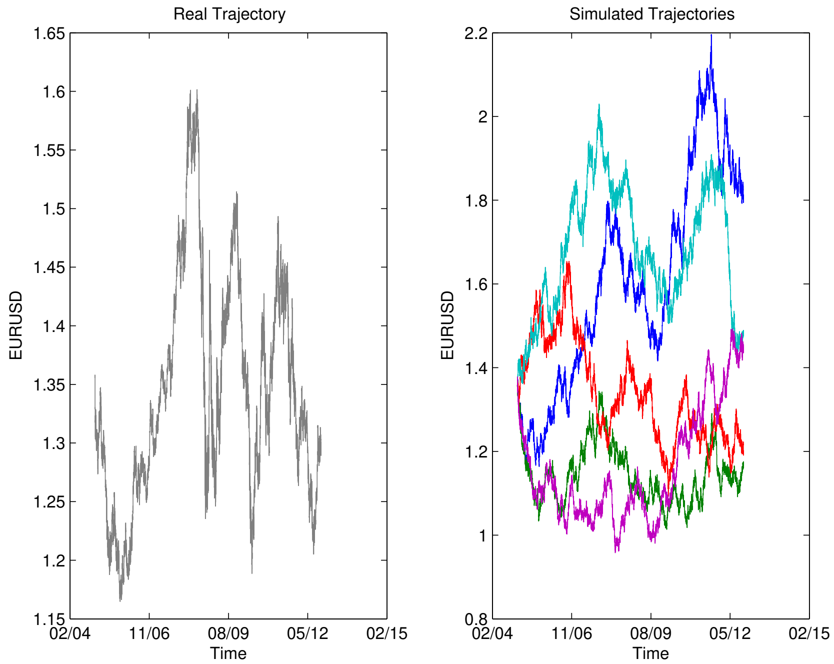

Figure 18: Real vs simulated trajectories – VG model; EURUSD

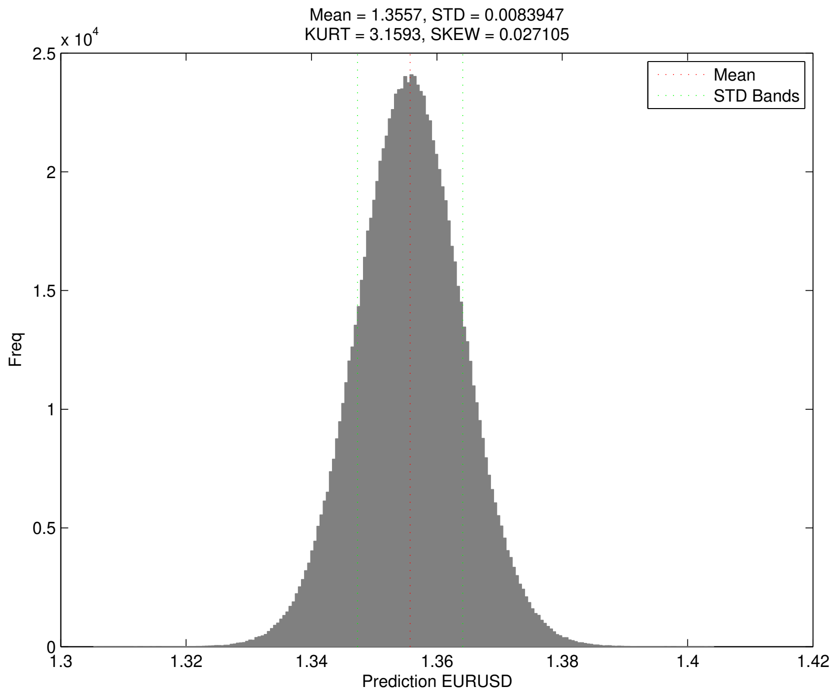

Figure 19: Simulated point prediction, 24h period ahead – VG model; fitted on EURUSD [1h returns 2005 – 2012]

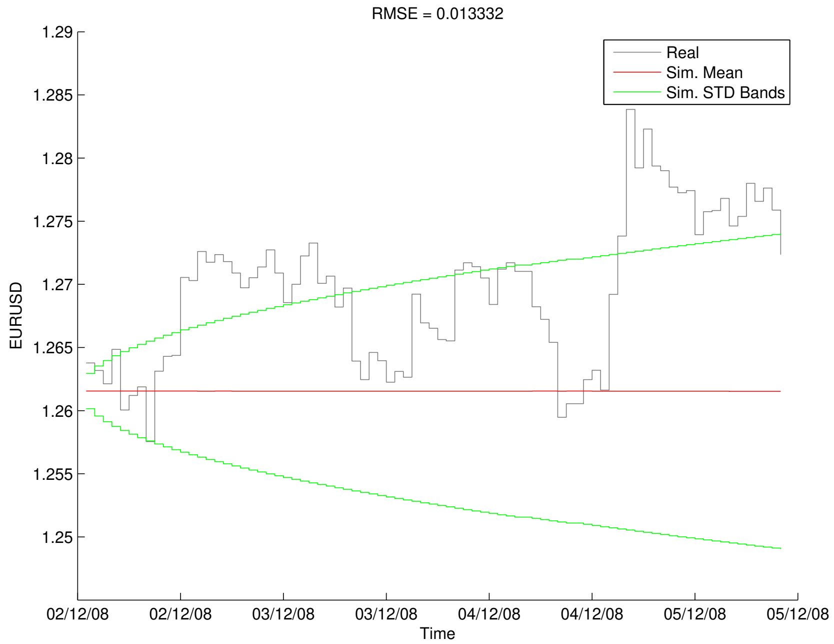

Figure 20: Out-of-sample simulation vs real trajectory – VG model; EURUSD

2.3.3 Meixner Model

Last example of (infinite activity) Lévy models we present is the *Meixner model* which was introduced to finance in [20]. Although there is a possibility to express the Meixner model process as a time-changed Brownian motion (Brownian subordination) – see [11] for instance – we do not present it here. The reason is that we use different approach for simulations, namely *rejection sampling*. This shall illustrate the fact that one may work with an infinite activity model even without usage of the convenient Brownian subordination concept.

Specification

Increments of the logarithmic FX rate (log-returns) are governed by the *Meixner distribution* $MXN(\mu\Delta t, \alpha, \beta, \delta\Delta t)$, hence density of the log-returns is written as follows

$$\left\{ \begin{array}{rcl} f_{\Delta r_t}(x) & = & \frac{(2\cos(\beta/2))^{2\delta\Delta t}}{2\alpha\pi\Gamma(2\delta\Delta t)} \exp\left(\frac{\beta(x-\mu\Delta t)}{\alpha}\right) \cdot \\ & \cdot & \left|\Gamma\left(\delta\Delta t + \frac{i(x-\mu\Delta t)}{\alpha}\right)\right|^2, \; x \in \mathbb{R}, \end{array} \right. \tag{8}$$

where Γ denotes the *gamma function*, $\alpha > 0$, $-\pi < \beta < \pi$, $\delta > 0$, and $\mu \in \mathbb{R}$. Note that μ is the drift (or location of returns) parameter, α controls scale of returns, β and δ determine shape of the distribution – skewness and kurtosis in particular.

Furthermore, we give *Lévy measure* of the Meixner model

$$\nu(\mathrm{d}z) = \frac{\delta \exp(\beta z/\alpha)}{z \sinh(\pi z/\alpha)} \mathrm{d}z \text{ on } \mathbb{R} \setminus \{0\}.$$

Estimation

We make use of MLE again. In this case, we plug (8) to (3) and perform maximization procedure in order to find optimal $\boldsymbol{\theta} = (\mu, \alpha, \beta, \delta)$. As initial values of parameters we use a MM estimate. Formulas for MM estimates are as follows

$$\hat{\delta} = (\hat{\mathcal{K}}_r - \hat{\mathcal{S}}_r^2 - 3)^{-1}, \quad \hat{\beta} = \mathrm{sign}(\hat{\mathcal{S}}_r) \arccos(2 - \hat{\delta}(\hat{\mathcal{K}}_r - 3)),$$
$$\hat{\alpha} = (\hat{\mathcal{V}}_r(\cos\hat{\beta} + 1))^{\frac{1}{3}}, \quad \hat{\mu} = \hat{\mathcal{M}}_r - \hat{\alpha}\hat{\delta}\tan(\hat{\beta}/2),$$

where $\hat{\mathcal{M}}_r$, $\hat{\mathcal{V}}_r$, $\hat{\mathcal{S}}_r$, and $\hat{\mathcal{K}}_r$ are *sample mean*, *sample variance*, *sample skewness*, and *sample kurtosis* (of the log-returns), respectively. Note that this MM estimate exists only if $\hat{\mathcal{K}}_r > 2\hat{\mathcal{S}}_r^2 + 3$. The estimate might be found in [13] for instance.

Statistics

Statistics of the Meixner model – namely *mean* $\mathcal{M}_r$, *variance* $\mathcal{V}_r$, *skewness* $\mathcal{S}_r$, and *kurtosis* $\mathcal{K}_r$ – are as follows

$$\mathcal{M}_r = (\mu + \alpha\delta\tan(\beta/2))\Delta t, \quad \mathcal{V}_r = \alpha^2\delta\Delta t/(\cos\beta + 1),$$
$$\mathcal{S}_r = \sin(\beta/2)\sqrt{\tfrac{2}{\delta\Delta t}}, \qquad \mathcal{K}_r = \tfrac{2-\cos\beta}{\delta\Delta t} + 3.$$

Returns Modeling

Mexiner model is the last model for which we present quality of fit Figure 21. Note that this model performs the best among all the models considered.

In this last case, we also give the simulation results. Model trajectories are illustrated in Figure 22. Point "prediction" distribution can be found in Figure 23. The last Figure 24 depicts mean and standard deviation bands of simulated trajectories.

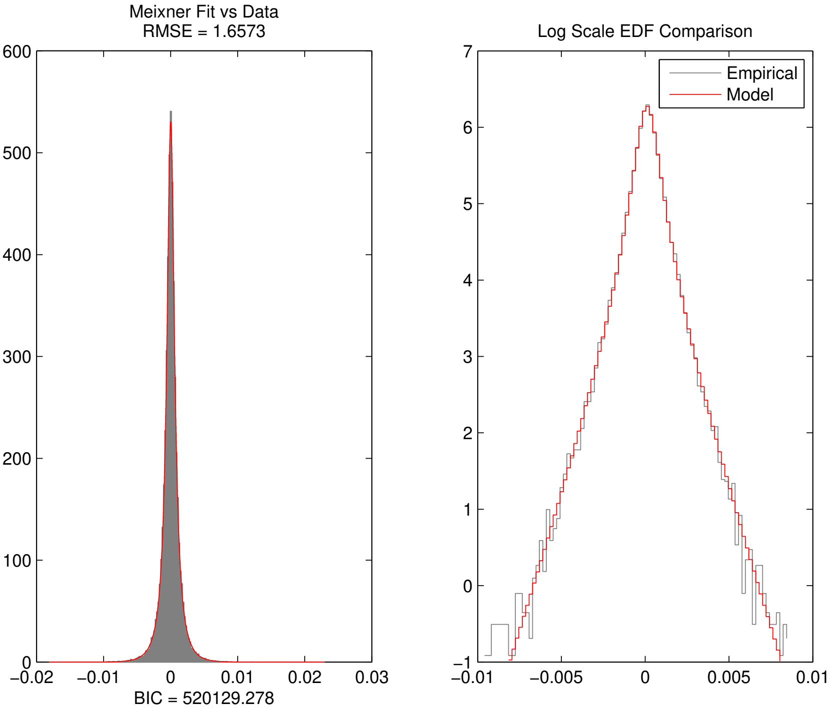

Figure 21: Comparison of empirical and model densities – Meixner model; EURUSD [1h returns 2005 – 2012]

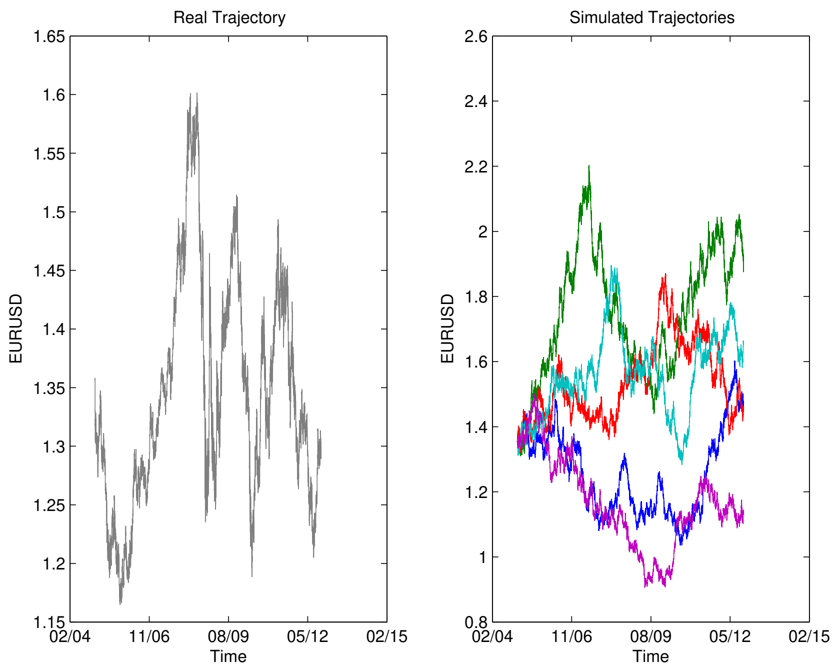

Figure 22: Real vs simulated trajectories – Meixner model; EURUSD

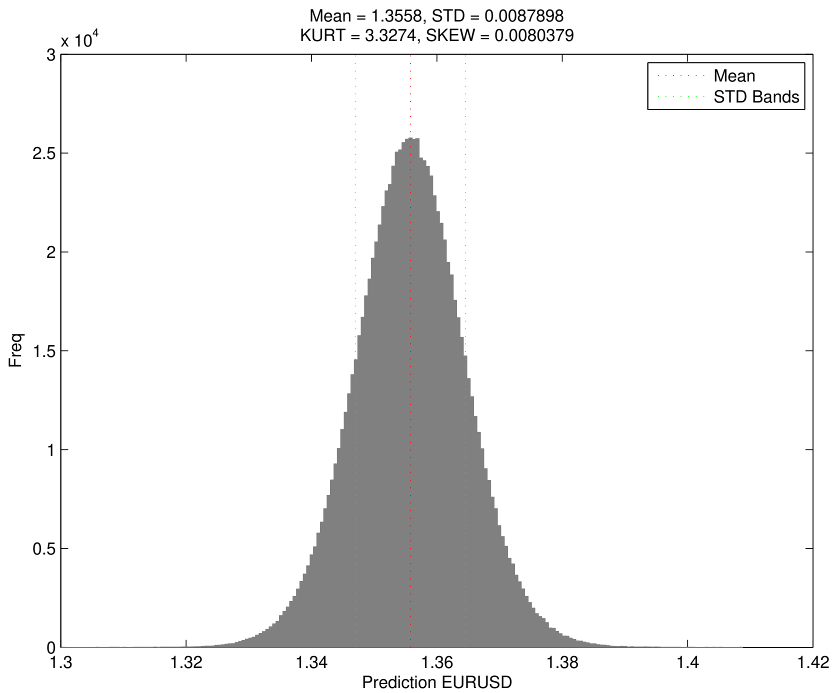

Figure 23: Simulated point prediction, 24h period ahead – Meixner model; fitted on EURUSD [1h returns 2005 – 2012]

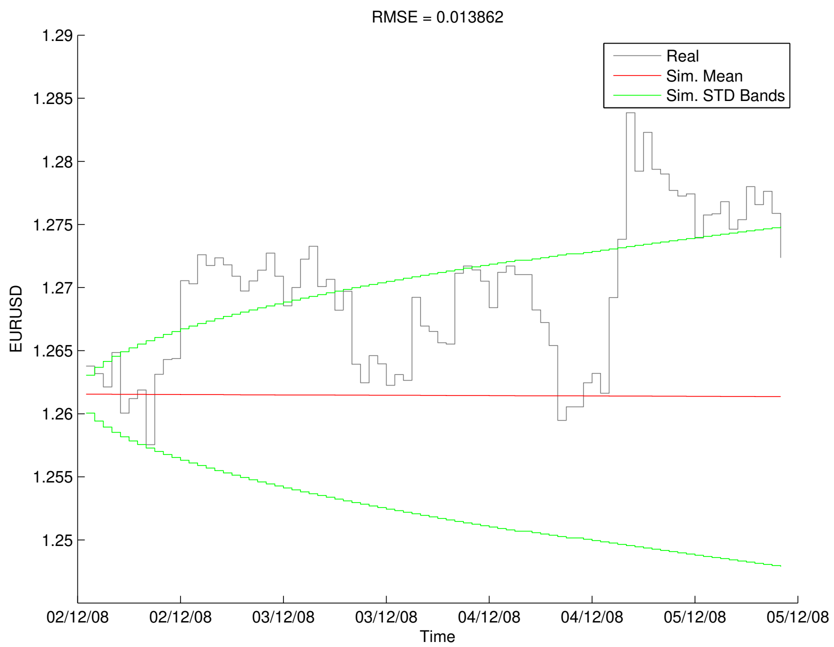

Figure 24: Out-of-sample simulation vs real trajectory – Meixner model; EURUSD

3 Concluding Remarks

We have considered multiple (jump) models in the context of FX rate modeling. We can support the evidence from the literature that jumps are important part of FX rate models. However, the focus is usually on the jump-diffusion models, although we have seen that some of the infinite activity models (especially the Meixner model) are quite capable of fitting the EURUSD returns in a proper way. This is probably due to the fact that it is easier to modify jump-diffusion models than an infinite activity models, since the structure of the former is more tractable (diffusion with a compound Poisson part included, say). As a possible suggestion of a future research, one may overcome this complication using the Brownian subordination concept.

As was shown by simulations, models in their standard forms do not differ a lot when it comes to a prediction. This has been attributed to the central limit theorem and the fact that in this Lévy (random walk) setting we always make sums of independent identically distributed random variables. However, this might change with a modification of standard forms of the models (as it is done to jump-diffusion models in the mentioned literature).

Of course, a similar composition may be conducted on different data and/or period; however, we believe that in this setting it provides an interesting set of information and a motivation to a further study of jump models role in the FX rates modeling. What might be also interesting is a comparison of (modified) jump models with models based on macroeconomic fundamentals (and possibly their combination).

References

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (9th printing)*. Dover Books, New York, 1972.

[2] Ole E. Barndorff-Nielsen. Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997.

[3] David S. Bates. Jumps and stochastic volatility: Exchange rate process implicit in Deutsche Mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.

[4] Thomas Busch, Bent Jasper Christensen, and Morten Ørregaard Nielsen. Forecasting exchange rate volatility in the presence of jumps. *Queen's Economics Department Working Paper* No. 1187, 2005.

[5] Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes.* Chapman & Hall / CRC Press, London, 2003.

[6] Jodie Duncan, John Randal, and Peter Thomson. Fitting jump diffusion using the EM algorithm. *Econometric Society Australasian Meeting*, 2009.

[7] Anders Eriksson, Lars Forsberg, and Eric Ghysels. Approximating the probability distribution of functions of random variables: A new approach. *CIRANO Working Papers* 2004s-21, 2004.

[8] George J. Jiang. Jump-diffusion model of exchange rate dynamics: Estimation via indirect inference. *University of Groningen*, 1998.

[9] Steven Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[10] Dilip B. Madan, Peter P. Carr, and Eric C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[11] Dilip B. Madan and Marc Yor. Representing the CGMY and Meixner Lévy processes as time changed Brownian motions. *The Journal of Computational Finance*, 12(1):27–47, 2008.

[12] John M. Maheu and Thomas H. McCurdy. Modeling foreign exchange rates with jumps. In David E. Rapach and Mark E. Wohar, editors, *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, pages 449–472. Emerald Group, 2008.

[13] Emanuele Mazzola and Pietro Muliere. Reviewing alternative characterizations of Meixner process. *Probability Surveys*, 8:127–154, 2011.

[14] José E. Figueroa, Steven R. Lancette, Kiseop Lee, and Yanhui Mi. Estimation of NIG and VG models for high frequency financial data. In Frederi G. Viens, Maria C. Mariani, and Ionut Florescu, editors, *Handbook of Modeling High-Frequency Data in Finance*, pages 3–26. John Wiley & Sons, 2012.

[15] Miloš Božović. The role of jumps in foreign exchange rates. *Universitat Pompeu Fabra Barcelona*, 2008.

[16] Robert C. Merton. Option pricing when underlying stocks are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[17] Ramzi Nekhili, Aslihan Altay-Salih, and Ramazan Gençay. Exploring exchange rates returns at different time horizons. *Physica A*, 313:671–682, 2002.

[18] Makoto Nirei and Vladyslav Sushko. Jumps in foreign exchange rates and stochastic unwinding of carry trades. *International Review of Economics & Finance*, 20(1):110–127, 2011.

[19] Bernt Øksendal and Agnès Sulem. *Applied Stochastic Control of Jump Diffusions.* Springer-Verlag, Berlin, 2005.

[20] Wim Schoutens and Jozef L. Teugels. Lévy processes, polynomials and martingales. *Communications in Statistics – Stochastic Models*, 14:335–349, 1998.