Stochastic differential equations applied to the study of geophysical and financial time series

Osei Kofi Tweneboah\textsuperscript{1}  
&  
Maria C. Mariani\textsuperscript{2}

\textsuperscript{1}Computational Science Program  
\textsuperscript{2}Department of Mathematical Sciences  
The University of Texas at El Paso

October 31, 2015
Outline

1. Introduction

2. The Model
   - The Gamma Process
   - Superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck Model

3. Estimation of Model Parameters

4. Model Simulation

5. Application of Model to Geophysical and Financial Time Series

6. Conclusions
Introduction

Recent research shows a renewed interest in describing ‘physical phenomena’ using new modeling techniques.

Long term correlations model have been applied to the occurrences of seismic events, [S. Lennartz et al., 2008].

Numerous deterministic and probabilistic models have been developed to depict various aspects of the mathematical modeling of seismic occurrence patterns and describing major stock indices.

In a recent work, scale invariant functions and Lévy models are applied to earthquake seismic data to estimate parameters related to some major earthquake events, [Mariani et al., 2013].
The procedure described in previous works seem to have a major drawback.

It is now widely accepted that the times of occurrence of a sequence of geophysical and financial time series on a given source may be stochastically dependent hence stochastic models may be more suitable.

Therefore in an attempt to overcome the modeling problems associated with the memory-less property models described in previous literature, we propose a continuous-time stationary and non-negative stochastic process that is useful in describing a unique type of dependence in a sequence of events.
We consider a continuous time stationary and non-negative process which is defined by the stochastic differential equation,

\[ dX_i(t) = -\lambda_i X_i(t) dt + dZ(\lambda_i t), \quad \lambda_i \in \mathbb{R}^+. \]  

(1)

where \( Z(\lambda_i t) \) is a subordinator i.e. a purely non-Gaussian Lévy process with positive increments (jumps).

The solution of Equation [1] is given by,

\[ X(t) = \sum_{i=1}^{m} w_i X_i e^{-\lambda_i t} + \int_0^t \sum_{i=1}^{m} w_i e^{-\lambda_i (t-s)} dZ(\lambda_i s) \]  

(2)

where \( \sum_{i=1}^{m} w_i = 1 \). The process \( X(t) \) is strictly positive.
The Gamma Process

A random variable $X$ has a gamma distribution $\Gamma(a, b)$ with rate and shape parameters, $a > 0$ and $b > 0$ respectively, if its density function is given by:

$$f_X(x; a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad \forall x > 0,$$

(3)

where $\Gamma$ denotes the gamma function.
The BDLP of Equation [2], \( Z(\lambda_i t)_{t \geq 0} \), is driven by a Gamma distribution with parameters \( a > 0 \) and \( b > 0 \) therefore Equation [2] is called superposed \( \Gamma(a, b) \) OU process.

The BDLP for the superposed \( \Gamma(a, b) \) OU process is a compound Poisson process.

The compound Poisson jumps a finite number of times in every compact time interval.

Therefore, the superposed \( \Gamma(a, b) \) OU process, \( X(t) \), jumps a finite number of times in a finite time interval.
A two component model of Equation [2] is given by,

\[
X(t) = w_1 X_1 e^{-\lambda_1 t} + \int_0^t w_1 e^{-\lambda_1 (t-s)} dZ(\lambda_1 s) \\
+ w_2 X_2 e^{-\lambda_2 t} + \int_0^t w_2 e^{-\lambda_2 (t-s)} dZ(\lambda_2 s), \quad t \geq 0.
\]

(4)

which, by a change of variable can also be written as

\[
X(t) = w_1 X_1 e^{-\lambda_1 t} + w_1 e^{-\lambda_1 t} \int_0^{\lambda_1 t} e^s dZ(s) \\
+ w_2 X_2 e^{-\lambda_2 t} + w_2 e^{-\lambda_2 t} \int_0^{\lambda_2 t} e^s dZ(s), \quad X_1, X_2 > 0, \lambda_1, \lambda_2 > 0.
\]

(5)
Estimation of Model Parameters

We aim at estimating the model parameters $a$ and $b$ from a sample of observations.

Proposition 1 relates the theoretical moments of $Z(1)$ with the theoretical moments of the stationary distribution of $\{X(t)\}_{t \geq 0}$.

**Proposition 1**

Suppose that $\{Z_t\}_{t \geq 0}$ is a Lévy process such that $E(Z(1)) = \mu < \infty$ and $Var(Z(1)) = \sigma^2 < \infty$. Assume that $\lambda_1, \lambda_2 > 0$, then the following are true.

1. $E(X_0) = \mu$.
2. $Var(X_0) = \frac{\sigma^2}{2}$.

From Proposition 1 the parameters $\mu$ and $\sigma^2$ relates to $a$ and $b$ as follows,

$$a = \frac{2\mu^2}{\sigma^2} \quad \text{and} \quad b = \frac{2\mu}{\sigma^2}. \quad (6)$$
Estimation of Model Parameters

We aim at estimating the intensity parameters $\lambda_2$ and $\lambda_1$ from a sample of observations.

The autocorrelation function of the process of Equation [5] is of the form,

$$\rho(\Delta t) = w_1 e^{-\lambda_1|\Delta t|} + w_2 e^{-\lambda_2|\Delta t|}, \forall \Delta t \in \mathbb{R}^+ \tag{7}$$

Assuming $\lambda_1 = \lambda_2$ in Equation [7], we have

$$\lambda_1 = -\frac{\log(\rho(\Delta t)) - \log(w_1 + w_2)}{\Delta t} \tag{8}$$

Wilog, we take $\Delta t = 1$. Therefore our estimate is

$$\lambda_1 = -\log(\rho(1)) + \log(w_1 + w_2) \tag{9}$$

where $\rho(1)$ denotes the empirical autocorrelation function at lag 1.
Model Simulation

We base our simulation on the solution of our proposed stochastic differential equation, Equation [5].

Simulation

To simulate Equation [5] in the time points \( t = n\Delta t, n = 0, 1, 2, \ldots \), we first simulate in the same time points as a Poisson process \( N = \{N_t, t \geq 0\} \) with intensity parameter \( a\lambda_1 \) and \( a\lambda_2 \), then,

\[
X_{n\Delta t} = w_1 X_{(n-1)\Delta t} e^{-\lambda_1 \Delta t} + w_1 \sum_{N_{(n-1)\Delta t+1}}^{N_{n\Delta t}} \chi_n \exp(-U_n \lambda \Delta t) \\
+ w_2 X_{(n-1)\Delta t} e^{-\lambda_2 \Delta t} + w_2 \sum_{N_{(n-1)\Delta t+1}}^{N_{n\Delta t}} \chi_n \exp(-U_n \lambda \Delta t)
\]

(10)

A Matlab module was developed to simulate Equation [10].
The superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model is applied to two time series arising in geophysics and finance. First of all, we estimate the model parameters $a, b$ and $\lambda_1$ using the relations given in Equations 6 and 9 respectively.
Next, we simulate the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model using the steps described in Equation 10. We simulated independent paths of our model using different time steps.

We compared the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model to the ordinary $\Gamma(a, b)$ Ornstein-Uhlenbeck model to check which of them best fits the data. In order to investigate our model fit, we computed the root mean square error for each region. The root mean square error indicates how well fitted is our model with respect to the given data set.
Background of the earthquake time series

- The geophysical data was obtained from U.S. Geological Survey (USGS) from January 1, 1973 to November 9, 2010.
- This data contains information about the date, longitude, latitude, and magnitude of each recorded earthquake in the region.
- The location of the major earthquake chosen defines the area studied.
- The location chosen cannot be too small due to lack of fit or too big because of noise from unrelated events.
- The earthquake magnitude is the recorded data used in the analysis.
Financial Time Series

Background of the financial time series

We studied emergent market indices corresponding to three countries:


The number of data points for BOVESPA, MERVAL and HSI is 2100, 1250 and 2675 respectively.
For a particular earthquake prone region, if the time series of data points (magnitude) of an earthquake are joined with lines then the following are the characteristics of the time series:

- Magnitude is a non-negative stationary process.
- For any finite interval of time there are only finite number of jumps.

The non-Gaussian superposed OU process has been shown to satisfy the properties of an earthquake time series. Thus making it a possible candidate for modeling earthquake time series.
For the geophysical data described above, results are presented for one randomly selected year where the earthquake magnitudes in different locations are available. We used data from 1973 for our analysis. Table 1 summarizes the results of the estimation of parameters for the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model.

**Table:** $\lambda_1, a, b.$

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of observations</th>
<th>$\lambda_1$</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973</td>
<td>653</td>
<td>2.4942</td>
<td>83.2756</td>
<td>18.6542</td>
</tr>
</tbody>
</table>
Table 2 and 3 summarize our numerical results for the geophysical time series when the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model and the ordinary $\Gamma(a, b)$ Ornstein-Uhlenbeck model were applied to real earthquake data series. We obtained $\lambda_2$ by adjusting $\lambda_1$ in order to fit the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model.

Table: Numerical results for the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973</td>
<td>2.4942</td>
<td>3.1942</td>
<td>0.50</td>
<td>0.50</td>
<td>0.1644</td>
</tr>
</tbody>
</table>

Table: Numerical results for the $\Gamma(a, b)$ Ornstein-Uhlenbeck model

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda_1$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973</td>
<td>2.4942</td>
<td>0.1664</td>
</tr>
</tbody>
</table>
Table 4 summarizes the results of the estimation of parameters for the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model.

Table: $\lambda_1, a, b$.

<table>
<thead>
<tr>
<th>Financial Indices</th>
<th>Number of Observations</th>
<th>$\lambda_1$</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOVESPA</td>
<td>2100</td>
<td>0.0017</td>
<td>5.4289</td>
<td>6.4362e-04</td>
</tr>
<tr>
<td>MERVAL</td>
<td>1250</td>
<td>0.2285</td>
<td>26.2184</td>
<td>0.0477</td>
</tr>
<tr>
<td>HSI</td>
<td>2674</td>
<td>0.0021</td>
<td>14.7389</td>
<td>0.0015</td>
</tr>
</tbody>
</table>

Table 5 and 6 summarizes the numerical results for each financial index when the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model and the ordinary $\Gamma(a, b)$ Ornstein-Uhlenbeck model were applied to time series respectively.
Table: Numerical results for the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model.

<table>
<thead>
<tr>
<th>Financial Indices</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOVESPA</td>
<td>0.0017</td>
<td>10</td>
<td>0.2</td>
<td>0.8</td>
<td>0.9098</td>
</tr>
<tr>
<td>MERVAL</td>
<td>0.2285</td>
<td>16</td>
<td>0.01</td>
<td>.99</td>
<td>0.2535</td>
</tr>
<tr>
<td>HSI</td>
<td>0.0021</td>
<td>6</td>
<td>0.1</td>
<td>0.9</td>
<td>0.2621</td>
</tr>
</tbody>
</table>

Table: Numerical results for the $\Gamma(a, b)$ Ornstein-Uhlenbeck model.

<table>
<thead>
<tr>
<th>Financial Indices</th>
<th>$\lambda_1$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOVESPA</td>
<td>0.0017</td>
<td>5.5643</td>
</tr>
<tr>
<td>MERVAL</td>
<td>0.2285</td>
<td>0.7668</td>
</tr>
<tr>
<td>HSI</td>
<td>0.0021</td>
<td>1.1970</td>
</tr>
</tbody>
</table>
Figures 1, 2 and 3 shows the sample path for the simulated data sets corresponding to BOVESPA, MERVAL and HSI respectively.

**Figure:** Sample Path of the BOVESPA financial index.
Analysis of Financial Time Series (Cont.)

Figure: Sample Path of the MERVAL financial index.

Figure: Sample Path of the HSI financial index.
Conclusions

- From the numerical results obtained, we conclude that in all instances the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model performed considerably well compared to the results for the ordinary $\Gamma(a, b)$ Ornstein-Uhlenbeck model. This is due to the fact that, the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model is a weighted sum of solutions.

- Moreover, because the superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck model need not be identically distributed, it offers a lot of flexibility in the model.

- Superposed $\Gamma(a, b)$ Ornstein-Uhlenbeck processes provide a class of continuous time processes which exhibits long memory behavior. The presence of long memory suggests that current information is highly correlated with past information at different levels, what may facilitate prediction.

- The methodology used in this work can be applied to other disciplines such as biology, bioinformatics, medicine and in social sciences.
Thank You!