UDK 519.234

APPLICATIONS OF SECOND ORDER ORNSTEIN UNLENBECK STOCHASTIC PROCESSES TO CREDIT RISK MODELING

M. Vaskouski
*Associate professor of
Belarusian State University,
Ph.D. in mathematics; senior
statistical analyst at HiQo
Solutions Ltd*

*Department of Higher Mathematics ,Belarusian State University, Republic of Belarus
HiQo Solutions, Ltd, USA
E-mail: vaskovskii@bsu.by, maxim.vaskouski@hiqo-solutions.com*

M. Vaskouski

Associate professor at Department of Higher Mathematics of Belarusian State University, Ph.D. in mathematics, senior statistical analyst at HiQo Solutions Ltd. I am an expert in theory of stochastic differential equations.

Abstract. We consider applications of second order stochastic processes for analysis and forecasting credit loss. In contrast to the Vasicek model based on the one-dimensional Ornstein-Uhlenbeck stochastic differential equation driven by the Wiener process, we study two-dimensional analogues of Ornstein-Uhlenbeck processes driven by fractional Brownian motions. These processes are applied to extrapolation of macroeconomic factors for modeling account loss probability. Second order Ornstein-Uhlenbeck stochastic processes capture local behavior of economic factors providing more realistic tools in comparison with the first order Ornstein-Uhlenbeck processes. The obtained results are applied to different types of account loss rate models in frame of FASB's Current Expected Credit Loss (CECL) and IASB's International Financial Reporting Standards 9 (IFRS 9) rules.

Keywords: Ornstein-Uhlenbeck processes, mean reverting, macroeconomic factors, rough path integration theory.

Introduction. Accounts loss rate models are core ingredients for credit loss estimation in frame of FASB's Current Expected Credit Loss (CECL) and IASB's International Financial Reporting Standards 9 (IFRS 9) rules [1]. Simple account loss models (time series, roll rate models) are based on fitting of loss rate time series by proper combinations of macroeconomic factors (e.g. unemployment rate, house price index, gross domestic product). More complicated models (age-period-cohort models, discrete time multihorizon survival models) extract exogenous components of historical accounts loss rate. These exogenous components, called environment functions, are modeled by macroeconomic factors.

Economic scenarios are applied to estimate future losses. According to the CECL rules, it is possible to use annual macroeconomic scenarios provided by the US Federal Reserve System for a short-term (24 months) economic prediction. Long-term economic prediction is based on using conception of mean reverting [1]. In contrast to the Vasicek model based on the one-dimensional Ornstein-Uhlenbeck stochastic differential equation driven by the Wiener process, we study two-dimensional analogues of Ornstein-Uhlenbeck processes driven by fractional Brownian motions.

Second order Ornstein-Uhlenbeck processes capture local behavior of economic factors and improve credit risk models in comparison with the first order Ornstein-Uhlenbeck processes.
This paper is organized in the following way. In section 2 we introduce necessary tools from rough path integration theory [2, 3]. In section 3 we deduce exact formulas for solutions of second order Ornstein-Uhlenbeck equations driven by fractional Brownian motions. In section 4 we derive new extrapolation method of macroeconomic factors by second order Ornstein-Uhlenbeck processes. Finally, Section 5 is devoted to application of the obtained results to CECL estimation in frame of the Age-Period-Cohort model [1, 4].

Elements of rough path integration theory

A stochastic process $B^H(t)$, $t \ge 0$, defined on a probability space (Ω, F, P) is called a fractional Brownian motion with Hurst index $H \in (0,1)$ if $B^H(t)$ is a Gaussian process with properties: 1) $B^H(t) = 0$ a.s., 2) $E(B^H(t)) = 0$ for any $t \ge 0$, 3) $E(B^H(t)B^H(s)) = \frac{1}{2}(t^{2H} + s^{2H} + |t-s|^{2H})$ for any $t, s \ge 0$ [5].

We consider a stochastic differential equation

$$dX(t) = b(X(t))dt + h(X(t))dW(t) + \sigma(X(t))dB^H(t),\ t \in [0,T], \qquad (1)$$

driven by independent Wiener process $W(t)$ and fractional Brownian motion with Hurst index $H \in \left(\frac{1}{3}, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right)$, where functions $b, h, \sigma : \mathbb{R} \to \mathbb{R}$ are deterministic.

Let us define the process $B(t) = (t, W(t), B^H(t))^T$ and function $B^H(t)$ $f(X) = (b(X), h(X), \sigma(X))$. We fix an arbitrary $\alpha \in \left(\frac{1}{3}, \min\left(\frac{1}{2}, H\right)\right)$ and consider α-Holder continuous rough path $\mathbf{B} = (B, \mathbb{B})$, where the two parametric random process $\mathbb{B} : [0,T]^2 \to \mathbb{R}^{3\times 3}$ is defined as follows:

$$\mathbb{B}^{(1,1)}(s,t) = \frac{1}{2}(t-s)^2,\quad \mathbb{B}^{(2,1)}(s,t) = \int_s^t (W(\tau) - W(s))d\tau,\quad \mathbb{B}^{(3,1)}(s,t) = \int_s^t (B^H(\tau) - B^H(s))d\tau,$$

$$\mathbb{B}^{(1,2)}(s,t) = -\mathbb{B}^{(2,1)}(s,t) + (t-s)(W(t) - W(s)),\quad \mathbb{B}^{(1,3)}(s,t) = -\mathbb{B}^{(3,1)}(s,t) + (t-s)(B^H(t) - B^H(s)),$$

$$\mathbb{B}^{(2,2)}(s,t) = \frac{1}{2}(W(t) - W(s))^2 - \frac{1}{2}(t-s),\quad \mathbb{B}^{(3,3)}(s,t) = \frac{1}{2}(B^H(t) - B^H(s))^2,$$

$$\mathbb{B}^{(2,3)}(s,t) \overset{L^2}{=} \lim_{|\Pi| \to 0} \left(W(t_k) - W(s)\right)\left(B^H(t_{k+1}) - B^H(t_k)\right),$$

$$\mathbb{B}^{(3,2)}(s,t) = -\mathbb{B}^{(2,3)}(s,t) + (W(t) - W(s))(B^H(t) - B^H(s)),$$

where $\Pi = \{s = t_0 < t_1 < \ldots < t_l = t\}$ is an arbitrary partition of segment $[s,t]$, $|\Pi| = \max_k (t_{k+1} - t_k)$.

Definition 1. A random process $X(t)$, $t \in [0,T]$, is called a solution of Eq. (1) if the process $X(t)$ is an F_t-adapted process (here F_t is the filtration generated by the processes $W(t)$, $B^H(t)$), admits α-Holder continuous paths a.s., and the function $t \to X(t,\omega)$ is a solution of the rough path differential equation $dz(t) = f(z(t))\mathbf{B}(t,\omega)$ for almost all $\omega \in \Omega$. Solution $X(t)$ of Eq. (1) with the initial condition $X(0)=x\in$ Ў is called unique if $P(Y(t)=Z(t)\forall t\in[0,T])=1$ for any two solutions $Y(t)$, $Z(t)$ of Eq. (1) with the initial condition $X(0)=x\in$ Ў.

Second order Ornstein-Uhlenbeck processes

Let us consider the following system of stochastic differential equations

$$\begin{cases} dx(t)=(\theta(\mu-x(t))+v(t))dt, \\ dv(t)=-\theta_1 v(t)dt+\sigma_1 dW(t)+\sigma_2 dB^H(t),\ t\in[0,T] \end{cases}, \tag{2}$$

where $W(t)$ and $B^H(t)$ are independent Wiener process and fractional Brownian motion with Hurst index $H\in\left(\frac{1}{3},\frac{1}{2}\right)\cup\left(\frac{1}{2},1\right)$; μ is a long term mean, σ_1 and σ_2 are non-negative volatility coefficients, θ and θ_1 are positive constants, $\theta\neq\theta_1$.

We consider the second equation of System (2) as a partial case of Eq. (1) with $B(t)=(t,W(t),B^H(t))^T$, $f(X)=(-\theta_1 X,\sigma_1,\sigma_2)$. Let us show that the process

$$v(t)=v_0e^{-\theta_1 t}+\int_0^t e^{-\theta_1(t-s)}dM(s) \tag{3}$$

is a solution of the second equation of System (2) with the initial condition $v(0)=v_0$, where integral in the right-hand side is the Young integral with respect to $M(t)=\sigma_1 W(t)+\sigma_2 B^H(t)$.

Since

$$v(t)=v(s)-\int_s^t \theta_1 v(\tau)d\tau+\sigma_1(W(t)-W(s))+\sigma_2(B^H(t)-B^H(s))$$

for any $t,s\in[0,T]$, the remainder terms can be estimated as follows:

$$R^{f(v)}(s,t):=f(v(t))-f(v(s))-Df(v(s))f(v(s))(B(t)-B(s))=O(|t-s|^{2\alpha}),$$

where $\alpha\in\left(\frac{1}{3},\min\left(\frac{1}{2},H\right)\right)$.

Taking into account the relation

$$\lim_{|\Pi|\to 0}\sum_{t_i,t_{i+1}\in\Pi} Df(v(t_i))f(v(t_i))\mathrm{B}(t_i,t_{i+1})=$$
$$=\lim_{|\Pi|\to 0}\sum_{t_i,t_{i+1}\in\Pi}(\theta_1^2 v(t_i)\Delta_i t-\theta_1\sigma_1\Delta_i W-\theta_1\sigma_2\Delta_i B^H)\Delta_i t=0,$$

where $\Delta_i g=g(t_{i+1})-g(t_i)$, we conclude that the rough path integral $\int_0^t f(v(s))d\mathbf{B}(s)$ coincides with the Young integral $\int_0^t f(v(s))dB(s)$. Therefore, relation (3) gives a solution of the second equation of System (2). Substituting this relation into System (2), we get

$$x(t) = \mu + C_1 e^{-\theta t} + C_2 e^{-\theta_1 t} + \frac{1}{\theta - \theta_1} \int_0^t (e^{-\theta_1 (t-s)} - e^{-\theta (t-s)}) dM(s), \tag{4}$$

where C_1 and C_2 are arbitrary constants. This family of processes is called second order Ornstein-Uhlenbeck processes.

Economic scenarios based on second order Ornstein-Uhlenbeck processes

Let $y(t)$, $t \in [t_{\min}, t_{\max}]$, be a time series of some macroeconomic factor. We are going to create extrapolation of time series $y(t)$ for $t > t_{\max}$ basing on the family of processes given by (4). It's clear that the best estimate of the process $x(t)$ is the conditional expectation $\bar{x}(t) = E(x(t) | F_0)$ with respect to the trivial σ-algebra F_0. Thus,

$$\bar{x}(t) = \mu + C_1 e^{-\theta t} + C_2 e^{-\theta_1 t}.$$

For convenience we suppose that the processes $y(t)$ and $\bar{x}(t)$ have the same discrete time scaling (usually we use one month as the minimal step that corresponds to $\Delta t = \frac{1}{12}$). Here μ is the long term mean (usually μ is selected to be equal to the average value across values of $y(t)$), parameters $C_1, C_2, \theta, \theta_1$ are free. Denote $t_n = t_{\max}, t_{n-1} = t_{\max} - \Delta t, y_n = y(t_n)$, $y_{n-1} = y(t_{n-1})$, $\bar{x}_n = \bar{x}(t_n)$, $\bar{x}_{n-1} = \bar{x}(t_{n-1})$. Setting up $y_n = \bar{x}_n$, $y_n - y_{n-1} = \bar{x}_n - \bar{x}_{n-1}$, we get values of C_1, C_2 which provide smooth continuation of the process $y(t)$ for $t > t_{\max}$:

$$C_2 = \frac{y_{n-1} - \mu - (y_n - \mu) e^{\Delta t}}{e^{-\theta_1 t_n} (e^{\theta_1} - e^{\theta})}, \quad C_1 = (y_n - \mu) e^{\theta t_n} - C_2 e^{(\theta - \theta_1) t_n}.$$

Without loss of generality we assume that $0 < \theta < \theta_1$. Then parameter θ corresponds to rate of reverting to mean value μ depending on the length of the economic cycle. Parameter θ_1 can be estimated from 24-months Consensus economic scenarios or FRB economic scenarios.

The confidence intervals corresponding to 95% probability can be estimated in the following way $x_{lower}(t) = \bar{x}(t) - 1.96\Xi(t)$, $x_{upper}(t) = \bar{x}(t) + 1.96\Xi(t)$, where

$$\Xi^2(t) = \frac{1}{(\theta - \theta_1)^2} E\left(\int_0^t (e^{-\theta_1 (t-s)} - e^{-\theta (t-s)}) dM(s) \right)^2.$$

Application to credit risk modeling

Let us consider discrete time Age-Period-Cohort model for estimation of Current Expected Credit Loss (CECL) on US mortgage loans, see [1, 4].

Denote by a, t, v age of credit (in months), calendar date and origination date, it's clear that $t = a + v$. We consider historical account loss rate (or Probability of Default) and attrition rate (or Probability of Attrition) defined as

$$PD(a,t) = \frac{N_{def}(a,t)}{N_{act}(a,t-1)}, \; PA(a,t) = \frac{N_{attr}(a,t)}{N_{act}(a,t-1)},$$

where N_{act} N_{act} , N_{def} N_{def} , N_{attr} N_{attr} are the numbers of active, charge-off, attrition accounts respectively.
Also we define exposure at default

$$EAD(a,t) = \frac{B_{def}(a,t)}{N_{def}(a,t)},$$

where B_{def} B_{def} is charge-off amount.
The first stage of analysis is to find Age-Period-Cohort decompositions for variables PD, PA, EAD:

$$\operatorname{logit}\left(PD(a,t)\right) = F_{PD}(a) + H_{PD}(t) + G_{PD}(v), \tag{5}$$

$$\operatorname{logit}\left(PA(a,t)\right) = F_{PA}(a) + H_{PA}(t) + G_{PA}(v), \tag{6}$$

$$\log\left(EAD(a,t)\right) = F_{EAD}(a) + H_{EAD}(t) + G_{EAD}(v), \tag{7}$$

here $\operatorname{logit}(p) = \log\left(\frac{p}{1-p}\right)$ $\operatorname{logit}(p) = \log\left(\frac{p}{1-p}\right)$ is the inverse function to the logistic function $g(z) = \frac{1}{1+e^{-z}}$ $g(z) = \frac{1}{1+e^{-z}}$ $z \in \mathbb{R}$.

The following graphs show the Age-Period-Cohort decomposition of account loss rate for US 30-years fixed rate mortgage loans with high risk (FICO Score is less than 660) [6]. This decomposition was obtained by method described in [7].

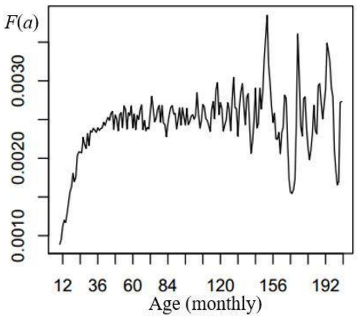

Figure *1a*. Lifecycle, PD, probability scale

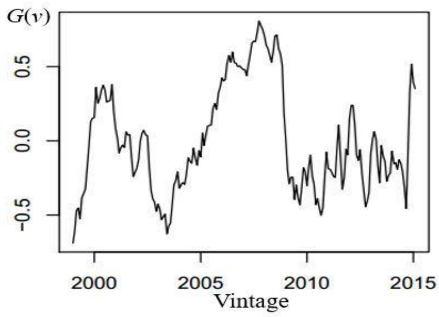

Figure *1b*. Credit Quality Function, PD

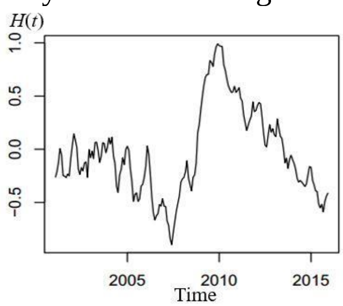

Figure *1c*. Environment Function, PD

The environment function $H_{PD}(t)$ $H(t)$ captures macroeconomic impact on account loss rate. We create a linear regression model (which is called economic model) for $H_{PD}(t)$:

$$H_{PD} \sim \beta_0 + \sum_{i=1}^{n} \beta_i f_i,$$

where $f_i(t)$ f_i ($i = 1, \dots, n$ $i = 1, \mathrm{K}, n$) are macroeconomic factors, β_i ($i = 1, \dots, n$ $i = 1, \mathrm{K}, n$) are regression coefficients. Usually a seasonality component is included as a part of the Age-Period-Cohort decomposition or economic model.

Table 1. – Parameters of the economic model for $H_{PD}(t)$ $H_{PD}(t)$.

Economic variable	Transformation	Optimal value of lag	Optimal value of win	Regression coefficient, β_i β_i
Real DPI	LogRatio	2	23	–1.734
Unemployment rate	Diff	5	24	0.078
House Price Index	LogRatio	12	17	–2.678
10-year Treasury yield	Diff	7	24	–0.106

Here the transformations LogRatio and Diff are defined as follows:

$$\mathrm{LogRatio}(x, t, lag, \mathrm{win}) = \log \frac{x(t - lag)}{x(t - lag - \mathrm{win})},$$

$$\mathrm{Diff}\,(x, t, lag, \mathrm{win}) = x(t - lag) - x(t - lag - \mathrm{win}).$$

Anyone, who wants to predict future loss, needs to know extrapolation of the environment function for forthcoming dates. Basing on economics history, we are able to calculate values of the economic fit $\hat{H}_{PD}$ up to present moment T_0. According to CECL rules, we get extrapolation of $\hat{H}_{PD}$ for the first two years by using FRB economic scenarios. Starting from $t = T_0 + 2$ we apply the economic model coefficients to the extrapolated macroeconomic factors, where mean reverting extrapolation based on second order Ornstein-Uhlenbeck processes is applied to the transformed economic factors (see Section 4).

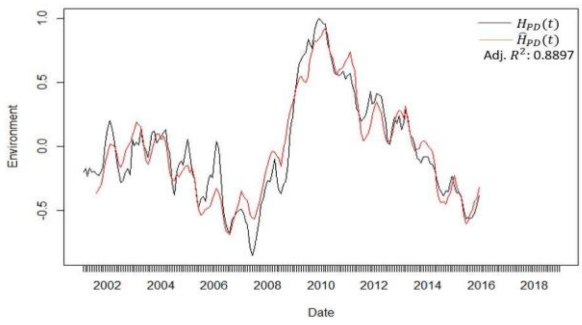

Figure 2.– Economic fit $\hat{H}_{PD}$ for the Environment Function $H_{PD}(t)$ H_{PD}

The same procedure is applied for the environment functions of the variables PA, EAD.

The lifecycle functions $F_{PD}(a)$ and $F_{PA}(a)$ have to be truncated after 10-years age and extrapolated with a constant. The lifecycle $F_{EAD}(a)$ should be truncated after 10-years age and extrapolated to be logarithmically decaying to zero.

Finally, CECL can be evaluated from relations (5) – (7), where the environment functions are replaced by extrapolated economic fits and original lifecycles are replaced by truncated and extrapolated lifecycles.

Remark 1. The same approach to get extrapolations of time functions can be applied to other credit risk models such as time series, roll rate, state transition, discrete time multihorizon survival models described in [1, 4].

Acknowledgment. The author thanks Dr J. Breeden and Dr A. Yablonski for their attention and valuable remarks.

References

[1.]Breeden, J.L. Living with CECL: Mortgage modeling alternatives / J.L. Breeden. – Middletown, 2018.

[2.]Levakov, A. Stochastic differential equations and inclusions / A. Levakov, M. Vaskouski. – Minsk: BSU, 2019.

[3.]Vaskouski, M. Asymptotic expansions of solutions of stochastic differential equations driven by multivariate fractional Brownian motions having Hurst indices greater than 1/3 / M. Vaskouski, I. Kachan // Stochastic Analysis and Applications. 2018. – Vol. 36, № 6. – P. 909–931.

[4.]Breeden, J.L. Current expected credit loss procyclicality: it depends on the model / J.L. Breeden, M. Vaskouski // Journal of Credit Risk. 2020. – Vol. 16, № 1. – P. 1 – 22. DOI: 10.21314/JCR.2020.258.

[5.]Mandelbrot, B.B. Fractional Brownian Motions, fractional noises and applications / B.B. Mandelbrot, J.W. van Ness // SIAM Review. 1968. – Vol. 10, № 4. – P. 422 – 437.

[6.]Vaskouski, M. Analysis and forecasting of credit loss with help of discrete time survival models / M. Vaskouski, A. Zadorozhnyuk // Proceedings of the International conference "Big Data and Advanced Analytics". 2019. – Vol. 1. – P. 285 – 293.

[7.]Schmid, V. Bayesian age-period-cohort modeling and prediction – BAMP/ V.J. Schmid, L. Held // Journal of Statistical Software. 2007. – Vol. 21, №8 – P. 1–15.