# Fitting a superposition of Ornstein–Uhlenbeck processes to time series of discharge in a perennial river environment

H. Yoshioka[1]

(Received 15 December 2021; revised 8 May 2022)

**Abstract**

Classical Ornstein–Uhlenbeck (OU) processes are Lévy-driven linear stochastic models with exponentially decaying autocorrelation functions which do not always fit more slowly decaying real time series data. A superposition of OU processes (known as a supOU process) is proposed to overcome this issue for application to river discharge time series data. The discharge data has a sub-exponential autocorrelation function and this is captured by the supOU process based on the mean reversion speed generated by a Gamma distribution. All the parameters of the supOU process are identified by matching the autocorrelation and the first to fourth statistical moments of the discharge data. The empirical and modelled histograms of the discharge data are comparable with each other.

DOI:10.21914/anziamj.v63.16985,  Published 2022-06-28, as part of the Proceedings of the 15th Biennial Engineering Mathematics and Applications Conference. ISSN 1445-8810. 

# Contents



# 1 Introduction

An Ornstein–Uhlenbeck (OU) process is a classical linear autoregressive model [13]. For the simplest 1D case it has a stationary probability density function (PDF), an autocorrelation function (ACF) which decays exponentially, and its statistical moments are known explicitly [11]. Owing to their convenience, OU processes have been widely used in scientific and engineering applications, including but not limited to insurance [6], environmental engineering [14], and flood analysis [10].

The exponential decay of the ACF of OU processes limits their usability because time series data encountered in applied problems often exhibit long memories, decaying only at algebraic rates [12]. Examples include hydrological processes, such as river discharge [4] and solute transport [8], which are key drivers of water environment and ecosystems. Nevertheless, modern hydrological models, including the author's work [14], still rely on OU processes, motivating us to extend the existing OU framework so that the statistics of interest can be efficiently evaluated and the ACFs are reasonably reproduced.

The objective of this article is to apply the superposition of OU processes, often referred to as a supOU process [1], to river discharge data as a case study. The supOU process is formally a superposition of infinitely many OU processes with different mean reversion speeds and mutually independent background driving Lévy processes with infinite activities. Applications of supOU processes are found in finance [5], physics [3] and related research areas, but not in hydrology and civil and environmental engineering, to the best of the author's knowledge.

Based on river discharge time series data, here it is shown that choosing a proper density to superpose OU processes enables us to reproduce the long memory of the ACF of the data. Moreover, its key statistical moments (average, standard deviation, skewness, kurtosis) are found explicitly. Based on these theoretical results, model parameter values corresponding to the real data are identified. Finally, it is shown that a plane Monte–Carlo method serves well for generating sample paths of the supOU process, despite its infinite-dimensional nature. This article contributes both analysis and a new engineering application to the study of supOU processes.

# 2 Mathematical model

The OU and supOU processes are reviewed, and statistical moments and the ACF of the latter are derived. Our explanation of the mathematical models follows the conventional ones [1].

## 2.1 The OU process

The classical 1D OU process is

$$dX_t = -\lambda(X_t - \underline{X})dt + dL_t\,, \quad t > 0\,. \tag{1}$$

Here, $t \in \mathbb{R}$ is the time, $X = (X_t)_{t\geqslant 0}$ is the OU process, $L = (L_t)_{t\geqslant 0}$ is a Lévy process, $\underline{X} \in \mathbb{R}$ is the reversion level (minimum discharge in our case), and $\lambda > 0$ is the reversion speed. The Lévy measure $\nu(dz)$ associated with $L$ is assumed to be a positive pure-jump subordinator; namely, it satisfies $\int_0^{+\infty} z\,\nu(dz) \in (0, +\infty)$. For the application example presented in Section 3 we assume $\nu(dz) = az^{-(\alpha+1)}e^{-bz}dz$ with parameters $a, b > 0$ and $\alpha < 1$.

Physically, equation (1) represents flood events of jumps and associated recessions. Generalising the results presented in this article to a generic case where the driving Lévy process is of a jump-diffusion type is straightforward.

Under a stationary state, equation (1) is solved as

$$X_t = \underline{X} + \int_{-\infty}^{t} e^{-\lambda(t-s)}\, dL_s\,, \quad t \in \mathbb{R}\,, \tag{2}$$

where the domain of the processes $X, L$ are extended to $\mathbb{R}$ [1]. A straightforward calculation shows that key statistics of $X$ at a stationary state are obtained analytically. Indeed, by denoting $\mathbb{E}$ as the expectation and setting $M_k = \int_0^{+\infty} z^k \nu(dz) \in (0, +\infty)$, $\mathbb{E}[X_t] = \underline{X} + M_1/\lambda$ and $\mathbb{E}[(X_t - \mathbb{E}[X_t])^2] = M_2/(2\lambda)$ are obtained. The representation (2) is a starting point for formulating a supOU process, as discussed in Section 2.2. Finally, the ACF with time lag $\tau \geqslant 0$ is

$$\text{ACF}(\tau) = \exp(-\lambda\tau)\,. \tag{3}$$

## 2.2 The supOU process

The supOU process at a stationary state is defined as

$$X_t = \underline{X} + \int_{-\infty}^{t} \int_{0}^{+\infty} e^{-\lambda(t-s)}\, \Lambda(d\lambda, ds)\,, \quad t \in \mathbb{R}\,, \tag{4}$$

where $\Lambda$ is an $\mathbb{R}$-valued pure-jump Lévy basis on $(0, +\infty) \times \mathbb{R}$ [1] with a probability measure $\pi$ of $\lambda$ on $(0, +\infty)$ and a background driving Lévy measure $\nu$. Because the reversion speed $\lambda$ physically corresponds to the inverse of the correlation time scale, the superposition made in the supOU process (the inner integration of (4) with respect to $\lambda$) implies the existence of multiple time scales in its sample paths. Intuitively, a random reversion speed is determined at each jump time.

The supOU process (4) is a semi-martingale and is expressed as a formal summation of infinitely many OU processes:

$$X_t = \sum_{i=1}^{\infty} \left( c_i \underline{X} + \int_{-\infty}^{t} \int_{0}^{+\infty} e^{-\lambda_i (t-s)} \, dL_s^{(i)} \right) = \sum_{i=1}^{\infty} X_t^{(i)}, \quad t \in \mathbb{R}, \tag{5}$$

with some non-negative series $\{c_i\}_{i=1,2,3,\dots}$ such that $\sum_{i=1}^{\infty} c_i = 1$, and where $X_t^{(i)}$ represents the OU process at $t$ with the reversion speed $\lambda_i$ and the Lévy process is $L^{(i)}$ with the measure $c_i \nu$ and each $L^{(i)}$ $(i = 1, 2, 3, \dots)$ is mutually independent. The OU process defined in Section 2.1 is a special case where the probability measure $\pi$ is a Dirac measure concentrated at a point. The well-posedness of the supOU process and its stationarity is satisfied if $R = \int_0^{+\infty} \lambda^{-1} \pi(d\lambda) \in (0, +\infty)$ [9, Theorem 2.2]. Only the average and variance have been derived in earlier research [9], but below we analytically find higher-order statistics.

At a stationary state the average (Ave), standard deviation (Std), skewness (Skew), and kurtosis (Kurt) of the supOU process are

$$\mathrm{Ave} = \mathbb{E}[X_t] = \underline{X} + RM_1\,, \tag{6}$$

$$\mathrm{Std} = \sqrt{\mathbb{E}[(X_t - \mathbb{E}[X_t])^2]} = \sqrt{\frac{RM_2}{2}}\,, \tag{7}$$

$$\mathrm{Skew} = \frac{\mathbb{E}[(X_t - \mathbb{E}[X_t])^3]}{\mathrm{Std}^3} = \frac{RM_3}{3\,\mathrm{Std}^3}\,, \tag{8}$$

$$\text{and} \quad \mathrm{Kurt} = \frac{\mathbb{E}[(X_t - \mathbb{E}[X_t])^4]}{\mathrm{Std}^4} - 3 = \frac{RM_4}{4\,\mathrm{Std}^4}\,. \tag{9}$$

The calculation procedure is straightforward as each expectation is calculated according to the definitions of each statistic and $\Lambda$, $\pi$, $\nu$. Finally, the ACF with time lag $\tau \geqslant 0$ is

$$\mathrm{ACF}(\tau) = \frac{\mathbb{E}[(X_t - \mathrm{Ave})(X_{t+\tau} - \mathrm{Ave})]}{\mathrm{Std}^2} = \frac{1}{R} \int_0^{+\infty} \frac{e^{-\lambda\tau}}{\lambda} \pi(d\lambda)\,. \tag{10}$$

This tractability of the supOU process is useful in engineering applications.

Assume the unimodal Gamma density

$$\pi(d\lambda) = \frac{1}{\Gamma(H)} \frac{\lambda^{H-1}}{B^H} \exp\left(-\frac{\lambda}{B}\right) d\lambda\,, \quad \lambda > 0\,, \tag{11}$$

with parameters $B > 0$ and $H > 1$, then $R = [B(H-1)]^{-1}$ follows and the ACF decays at an algebraic rate:

$$\text{ACF}(\tau) = \left(\frac{1}{1 + B\tau}\right)^{H-1}. \tag{12}$$

This contrasts with the exponential form (3) of the OU process and suggests that appropriately randomising the reversion speed leads to a long memory where multiple time scales coexist.

# 3 Application to real data

The supOU process is fitted against time series data of river discharge from a perennial river environment in a mountainous region. The study site is Takeno-hana station at Tabusa River, a branch of the Gono River (a first-class river in Japan) in Hiroshima Prefecture, Japan, for which the hourly record of river discharge since January 2017 is publicly available.[1] This river is a habitat of inland fishery resources, including *Plecoglossus altivelis altivelis* (Ayu), and its streamflow environment and habitat quality have recently been of research interest.

Figure 1 shows the four-year discharge time series data (from January 2017 to December 2020) of Takeno-hana station and Figure 2 presents its ACF. Figure 1 suggests that the discharge time series is driven by jumps decaying toward the minimum value $\underline{X}$, which has been determined from the data as $0.1\,\text{m}^3/\text{s}$. The black symbols in Figure 2 show the empirical ACF and its

[1]Water Information System http://www1.river.go.jp/cgi-bin/SiteInfo.exe?ID=307051287713030.

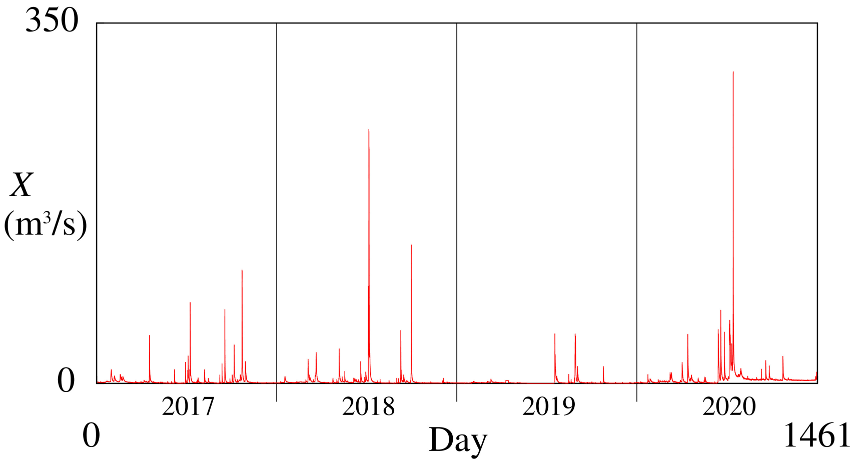


Figure 1: The hourly discharge time series of Takeno-hana station.

algebraic fit (12) for which a least-squares method gives $B = 0.108$ per hour and $H = 1.75$, where ensuring the empirical ACF is positive is utilised in the fitting. The fitting results show that the ACF of the discharge time series has an algebraic decay, and furthermore it has a long memory [9] because

$$\int_0^{+\infty} \mathrm{ACF}(\tau)\, d\tau = +\infty\,, \tag{13}$$

which is qualitatively different from the exponential fit (3) for which the integral remains finite. The fitting of Figure 2 suggests an advantage of the supOU process over the classical OU process in the present case.

Having obtained the values of $B$ and $H$, the Lévy measure $\nu$ is identified. The theoretical Ave, Std, Skew, and Kurt are fitted against the empirical ones by a nonlinear least-squares method so that the following objective is minimised:

$$\left(\frac{\mathrm{Ave}_m - \mathrm{Ave}_e}{\mathrm{Ave}_e}\right)^2 + \left(\frac{\mathrm{Std}_m - \mathrm{Std}_e}{\mathrm{Std}_e}\right)^2 +$$

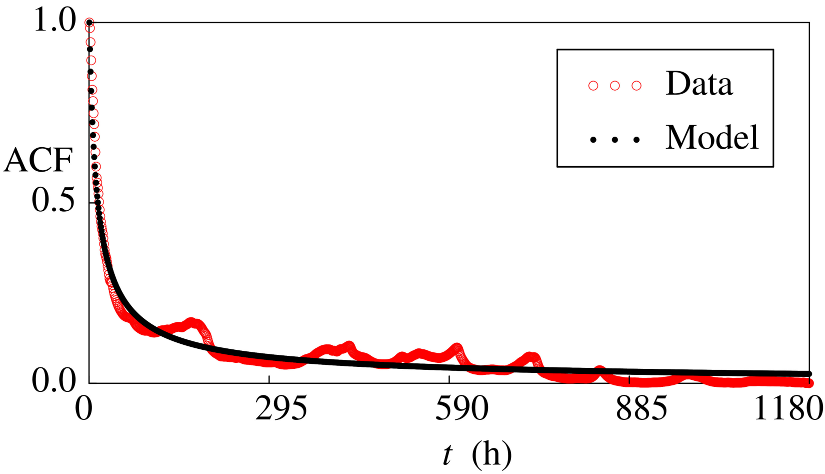


Figure 2: The empirical (red) and fitted (black) ACFs.

$$\left(\frac{\text{Skew}_m - \text{Skew}_e}{\text{Skew}_e}\right)^2 + \left(\frac{\text{Kurt}_m - \text{Kurt}_e}{\text{Kurt}_e}\right)^2 . \qquad (14)$$

Here, the subscripts $e$ and $m$ stand for 'empirical' and 'modelled', respectively, and the parameters to be optimised are $a$, $b$ and $\alpha$ of the Lévy measure $\nu$. An analogous fitting method has been successfully applied to an OU process [14].

Table 1 compares the empirical and fitted statistics and their relative errors with the best fit values $\alpha = 0.596$, $a = 0.0130\,\text{m}^{3\alpha}/\text{s}^{\alpha}$ per hour, and $b = 0.00806\,\text{s}/\text{m}^3$, demonstrating that the assumed model can reproduce the statistics within the relative errors of several percent. The result suggests that the discharge time series involves infinitely many small fluctuations as the background driving Lévy process satisfies $\alpha > 0$. The relative errors are even smaller than $1\%$ for Ave and Std.

Finally, the performance of the model with the identified parameter values is examined by comparing probability density functions (PDFs). Because the PDF of the supOU process is not available analytically, a Monte–Carlo method

Table 1: Comparison of the empirical and modelled statistics and their relative errors.

| Statistics | Empirical | Model | Relative error |
|---|---|---|---|
| Ave ($m^3/s$) | $2.59 \times 10^0$ | $2.58 \times 10^0$ | $6.23 \times 10^{-3}$ |
| Std ($m^3/s$) | $7.84 \times 10^0$ | $7.88 \times 10^0$ | $5.92 \times 10^{-3}$ |
| Skew (–) | $1.61 \times 10^1$ | $1.47 \times 10^1$ | $8.38 \times 10^{-2}$ |
| Kurt (–) | $4.05 \times 10^2$ | $4.19 \times 10^2$ | $3.42 \times 10^{-2}$ |

is employed. By exploiting the randomisation algorithm for supOU processes driven by compound Poisson processes [9], the following discretisation is proposed to generate sample paths:

$$X_{(n+1)\Delta t} = e^{-\lambda_n \Delta t} X_{n\Delta t} + (1 - e^{-\lambda_n \Delta t})\underline{X} + \frac{1 - e^{-\lambda_n \Delta t}}{\lambda_n \Delta t} \Delta L_n \,, \quad n = 0, 1, 2, \ldots \,, \tag{15}$$

with an initial condition $X_0 > 0$. Here, $\Delta t = 0.001$ hours is the time increment generated by the rejection sampling method [7], each $\Delta L_n$ is a mutually independent discretised tempered stable process, and $\lambda_n$ is a random variable generated by the identified density proportional to $\pi/\lambda$ with each $\lambda_n$ being mutually independent.

With the initial condition $X_0 = \underline{X}$, a long sample path with length 10 000 years was generated for the PDF to be numerically evaluated. The statistics computed by the Monte–Carlo method are Ave $= 2.57\,m^3/s$, Std $= 7.87\,m^3/s$, Skew $= 14.8$, and Kurt $= 422$, which are close to the theoretical values in Table 1. Figure 3 compares the empirical and modelled PDFs, showing that the proposed model serves well from the viewpoint of the PDF, although the empirical PDF is slightly sharper than the modelled one. In particular, both the empirical and modelled results consistently show the convex and decreasing PDFs.

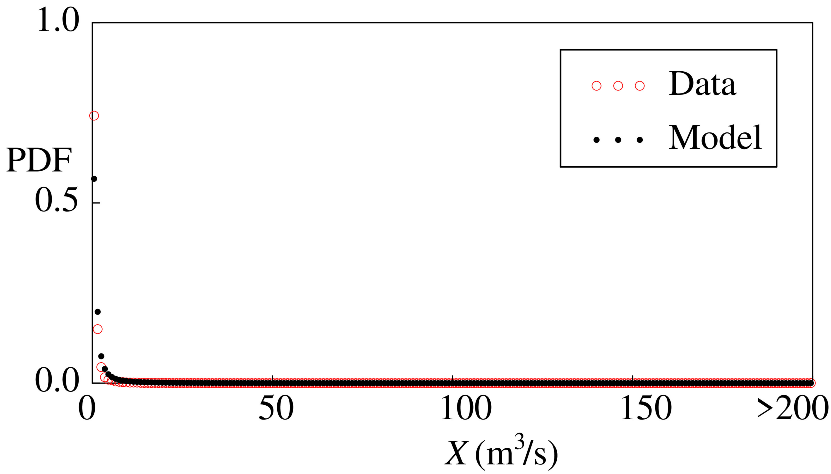


Figure 3: The empirical (red) and fitted (black) probability density functions (PDFs).

# 4 Conclusion

A supOU process was fitted to time series data of a four-year river discharge. Theoretically, the proposed identification method can be applied to all the perennial rivers, while its performance would be different for different rivers. The author is currently studying the applicability of supOU processes and related mathematical models in the analysis of riverine hydrological processes, including Volterra processes as representative non-Markovian models [2]. Detailed convergence analysis of the Monte–Carlo method is also an important issue to be addressed in the future. A research project to optimally control supOU processes is currently underway. In this case, the control problem will be reduced to solving a dynamic programming equation in an infinite-dimensional space. The author is currently investigating an infinite-dimensional Riccati equation for regulating discharge.

**Acknowledgements** This research was supported by JSPS Grant 19H03073, the Kurita Water and Environment Foundation (Grants 19B018 and 20K004), a Grant for Environmental Research Projects from the Sumitomo Foundation (203160), and Grant B4R31001 from MLIT, Japan. The hydrological data from Takeno-hana station in 2020 was provided courtesy of MLIT Haidzuka Dam branch.

# References

[1] O. E. Barndorff-Nielsen. "Superposition of Ornstein–Uhlenbeck type processes". In: *Theory Prob. Appl.* 45.2 (2001), pp. 175–194. DOI: 10.1137/S0040585X97978166 (cit. on pp. C85, C86, C87).

[2] O. E. Barndorff-Nielsen, F. E. Benth, and A. E. D. Veraart. "Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes". In: *Bernoulli* 19.3 (2013), pp. 803–845. DOI: 10.3150/12-BEJ476 (cit. on p. C93).

[3] O. E. Barndorff-Nielsen and N. N. Leonenko. "Burgers' turbulence problem with linear or quadratic external potential". In: *J. Appl. Prob.* 42.2 (2001), pp. 550–565. URL: http://www.jstor.org/stable/30040809 (cit. on p. C86).

[4] J. Beran, Y. Feng, S. Ghosh, and R. Kulik. *Long-Memory Processes.* Springer-Verlag, Berlin, Heidelberg, 2016. DOI: 10.1007/978-3-642-35512-7 (cit. on p. C85).

[5] F. Fuchs and R. Stelzer. "Mixing conditions for multivariate infinitely divisible processes with an application to mixed moving averages and the supOU stochastic volatility model". In: *ESAIM: Prob. Stat.* 17 (2013), pp. 455–471. DOI: 10.1051/ps/2011158 (cit. on p. C86).

[6] Y. Kabanov and S. Pergamenshchikov. "Ruin probabilities for a Lévy-driven generalised Ornstein–Uhlenbeck process". In: *Fin. Stoch.* 24.1 (2020), pp. 39–69. DOI: 10.1007/s00780-019-00413-3 (cit. on p. C85).

[7] R. Kawai and H. Masuda. "On simulation of tempered stable random variates". In: *J. Comput. Appl. Math.* 235.8 (2011), pp. 2873–2887. DOI: 10.1016/j.cam.2010.12.014 (cit. on p. C92).

[8] S. Pelacani and F. G. Schmitt. "Scaling properties of the turbidity and streamflow time series at two different locations of an intra-Apennine stream: Case study". In: *J. Hydro.* 603.B (2021), p. 126943. DOI: 10.1016/j.jhydrol.2021.126943 (cit. on p. C85).

[9] R. Stelzer, T. Tosstorff, and M. Wittlinger. "Moment based estimation of supOU processes and a related stochastic volatility model". In: *Stat. Risk Model.* 32.1 (2015), pp. 1–24. DOI: 10.1515/strm-2012-1152 (cit. on pp. C88, C90, C92).

[10] S. Suweis, E. Bertuzzo, G. Botter, A. Porporato, I. Rodriguez-Iturbe, and A. Rinaldo. "Impact of stochastic fluctuations in storage-discharge relations on streamflow distributions". In: *Water Resource. Res.* 46.3 (2010), W03517. DOI: 10.1029/2009WR008038 (cit. on p. C85).

[11] M. Tamborrino and P. Lansky. "Shot noise, weak convergence and diffusion approximations". In: *Physica D: Nonlinear Phenomena* 418 (2021), p. 132845. DOI: 10.1016/j.physd.2021.132845 (cit. on p. C85).

[12] E. Taufer and N. Leonenko. "Simulation of Lévy-driven Ornstein–Uhlenbeck processes with given marginal distribution". In: *Comput. Stat. Data Anal.* 53.6 (2009), pp. 2427–2437. DOI: 10.1016/j.csda.2008.02.026 (cit. on p. C85).

[13] C. Van Den Broeck. "On the relation between white shot noise, Gaussian white noise, and the dichotomic Markov process". In: *J. Stat. Phys.* 31 (1983), pp. 467–483. DOI: 10.1007/BF01019494 (cit. on p. C85).

[14] H. Yoshioka and Y. Yoshioka. "Designing cost-efficient inspection schemes for stochastic streamflow environment using an effective Hamiltonian approach". In: *Opt. Eng.* (2021), pp. 1–33. DOI: 10.1007/s11081-021-09655-7 (cit. on pp. C85, C91).

## Author address

1. **H. Yoshioka**, Graduate School of Natural Science and Technology, Shimane University, Nishikawatsu-cho 1060, Matsue, Japan.
   mailto:yoshih@life.shimane-u.ac.jp
   orcid:0000-0002-5293-3246