mathematics

Article

Modeling Dynamic Regime Shifts in Diffusion Processes: Approximate Maximum Likelihood Estimation for Two-Threshold Ornstein–Uhlenbeck Models

Svajone Bekesiene [1,*], Anatolii Nikitin [2,3] and Serhii Nechyporuk [4]

1 Research Group on Logistics and Defence Technology Management, General Jonas Zemaitis Military Academy of Lithuania, Silo 5a, 10322 Vilnius, Lithuania
2 Educational and Scientific Institute of Information Technology and Business, The National University of Ostroh Academy, UA-03580 Ostroh, Ukraine; anatolii.nikitin@oa.edu.ua
3 Faculty of Natural Sciences, Jan Kochanowski University, Stefana Zeromskiego 5, 25-369 Kielce, Poland
4 Department of Economics, The National University of Ostroh Academy, UA-03580 Ostroh, Ukraine; serhii.a.nechyporuk@oa.edu.ua
* Correspondence: svajone.bekesiene@lka.lt

Abstract

This study addresses the problem of estimating parameters in a two-threshold Ornstein–Uhlenbeck diffusion process, a model suitable for describing systems that exhibit changes in dynamics when crossing specific boundaries. Such behavior is often observed in real economic and physical processes. The main objective is to develop and evaluate a method for accurately identifying key parameters, including the threshold levels, drift changes, and diffusion coefficient, within this stochastic framework. The paper proposes an iterative algorithm based on approximate maximum likelihood estimation, which recalculates parameter values step by step until convergence is achieved. This procedure simultaneously estimates both the threshold positions and the associated process parameters, allowing it to adapt effectively to structural changes in the data. Unlike previously studied single-threshold systems, two-threshold models are more natural and offer improved applicability. The method is implemented through custom programming and tested using synthetically generated data to assess its precision and reliability. The novelty of this study lies in extending the approximate maximum likelihood framework to a two-threshold Ornstein–Uhlenbeck process and in developing an iterative estimation procedure capable of jointly recovering both threshold locations and regime-specific parameters with proven convergence properties. Results show that the algorithm successfully captures changes in the process dynamics and provides consistent parameter estimates across different scenarios. The proposed approach offers a practical tool for analyzing systems influenced by shifting regimes and contributes to a better understanding of dynamic processes in various applied fields.

Keywords: approximate maximum likelihood method; two-threshold diffusion process; stochastic differential equation

MSC: 37M25; 60H10

check for updates

Academic Editor: Luigi Rodino

Received: 12 September 2025
Revised: 25 October 2025
Accepted: 27 October 2025
Published: 29 October 2025

Citation: Bekesiene, S.; Nikitin, A.; Nechyporuk, S. Modeling Dynamic Regime Shifts in Diffusion Processes: Approximate Maximum Likelihood Estimation for Two-Threshold Ornstein–Uhlenbeck Models. *Mathematics* **2025**, *13*, 3450. https://doi.org/10.3390/math13213450

1. Introduction

The threshold effect refers to a phenomenon in which changes in an outcome remain minimal or nonexistent while a variable changes up to a certain point. Once this threshold is exceeded, the effect becomes abrupt or highly pronounced. Numerous studies have focused on developing and analyzing stochastic models that incorporate threshold effects, for example [1,2]. Modeling systems with multiple threshold effects also represents an important practical challenge. For instance, in finance, small increases in income can trigger a sharp change in the effective tax rate once a threshold is crossed [3]. In clinical pharmacology, the therapeutic efficacy of a drug exhibits a concentration-dependent profile, in which sub-therapeutic or supra-therapeutic blood concentrations can result in diminished effectiveness or adverse effects, respectively [4]. Here, there are both minimum and maximum safe thresholds. Unlike single-threshold models, two-threshold models more naturally reflect such real-world processes and allow for greater flexibility in capturing regime dynamics [5,6].

Similarly, the Ornstein–Uhlenbeck process considered in this article is used to describe phenomena that tend to fluctuate around an equilibrium state under the influence of random noise [7]. For instance, it is commonly applied to the modeling of interest rates, asset prices, and other financial variables [8]. Threshold diffusion processes of this type enable the analysis of distinct dynamic regimes while accounting for the current state of the system [9–11].

Gaussian-type threshold diffusion models that incorporate location shifts have been investigated in [8,12–14]. Parameter estimation techniques for single-threshold models were examined in [9,12]. A quasi-probabilistic method for estimating shift parameters in two-regime threshold diffusion processes was proposed in [15]. In [13], an initial testing procedure based on an approximate maximum likelihood framework was developed to assess linearity of the diffusion process and the presence of threshold effects in both drift and diffusion components. Convergence properties, which constitute a key aspect of the asymptotic behavior of the approximate maximum likelihood estimator, were analyzed in [16–18].

Unlike Bayesian methods, which rely on simulating parameter distributions through random sampling and require prior assumptions about their values [19], the AMLE approach provides direct parameter estimation based on the likelihood function without using such prior hypotheses, which significantly simplifies computations.

Compared to the Expectation–Maximization (EM) algorithm [20], which iteratively maximizes conditional expectations and may exhibit slow convergence in complex or flat regions of the parameter space, the proposed iterative AMLE procedure employs explicit numerical updates with a convergence check based on the spectral radius condition $p(J) > 1$.

This study investigates a two-threshold, three-mode Ornstein–Uhlenbeck diffusion process. The purpose of the study is to develop an approximate log-likelihood function for a process of this type, as well as to estimate the parameters of each of the modes and determine the threshold values, and to verify the proposed methodological approach using software code.

We employed a comprehensive set of methods to achieve this goal. The process was discretized by the Euler scheme, and a likelihood function was formed. The set of parameters for each mode is calculated by solving a system of equations with coefficients that take into account the variability of the data. The threshold values are calculated using an iterative algorithm in which, at each step the mode boundaries are checked for convergence.

The remainder of the paper is structured as follows: Section 2 outlines the mathematical background and discusses existing approaches to modeling systems with threshold behavior; Section 3 introduces the proposed iterative algorithm for estimating model parameters and threshold values; Section 4 details the iterative estimation procedure and explains the convergence criteria; Section 5 presents the numerical implementation, describes the simulation setup, and reports the main results; and Section 6 concludes the paper with a summary of findings and potential directions for future research.

2. Mathematical Background and Approaches to Research

The Ornstein–Uhlenbeck process [21], which has a number of properties (stationarity, Gaussianity, Markov property), is defined from the stochastic differential Equation (1):

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, X_0 = X(t_0), \tag{1}$$

where X_t is the state variable at time t; $\theta > 0$ is mean-reversion rate, indicating how quickly the process tends to return to the long-term mean μ; $\sigma > 0$ is the diffusion coefficient, representing the intensity of the random fluctuations; W_t is a Wiener process (or standard Brownian motion), $\mathbf{E}[W_t] = 0$, $\mathrm{Var}(W_t) = t$.

Equation (1) has an analytical solution at $t_0 = 0$ [9]:

$$X_t = X_0 e^{-\theta t} + \mu\left(1 - e^{-\theta t}\right) + \sigma \int_0^t e^{-\theta(t-s)}\, dW_s.$$

Unlike the single-threshold systems commonly examined in the literature [10], the present study advances the methodological framework by introducing two distinct threshold values into the stochastic process. This extension facilitates the representation of systems governed by three operational regimes, thereby improving the model's ability to capture more complex and realistic dynamics. Two-threshold models are especially useful in applications where structural shifts occur in a non-symmetric or multi-phase manner, such as in economics, climatology, and biological systems [5,22]. By incorporating an intermediate state between two extreme regimes, this approach offers a more natural and flexible characterization of real-world processes in which transitions across multiple dynamic states frequently arise.

Accordingly, the stochastic process with two modes and two thresholds can be mathematically formulated as follows in Equation (2):

$$\begin{aligned} dX_t &= \{(\beta_{10} + \beta_{11}X_t)I(X_t \le r_1) + (\beta_{20} + \beta_{21}X_t)I(r_2 > X_t \ge r_1) \\ &\quad + (\beta_{30} + \beta_{31}X_t)I(X_t \ge r_2)\}dt \\ &\quad + \{\sigma_1 I(X_t \le r_1) + \sigma_2 I(r_2 > X_t \ge r_1) + \sigma_3 I(X_t > r_2)\}dW_t, \end{aligned} \tag{2}$$

where β_{i0}, β_{i1},—determine the linear shift in the i-th mode ($i = 1, 2, 3$), and σ_i—the diffusion coefficient in this mode. The process is divided by thresholds r_1 and r_2 into three modes and each mode has different shift and diffusion coefficients. The behavior of the Ornstein–Uhlenbeck process with two thresholds changes depending on which of the three modes X_t is in.

In applied problems, the analysis of a stochastic process is carried out at discrete points in time, so we will consider a discretized version of the process. Let $\{X_0, X_1, \ldots, X_q\}$ be a set of data of the process X_t, which were obtained at times $\{\{t_0, t_1, \ldots, t_q\}$, where $\Delta_j = t_j - t_{j-1}$, is the time step between the obtained values.

To discretize stochastic differential equations, we will use the Euler scheme (3):

$$\begin{gathered} X_j - X_{j-1} = \\ = \Delta_j\Big\{(\beta_{10} + \beta_{11}X_{j-1})I(X_{j-1} \le r_1) + (\beta_{20} + \beta_{21}X_{j-1})I\Big(r_2 > X_{j-1} \ge r_1\Big) + (\beta_{30} + \beta_{31}X_{j-1})I(X_{j-1} \ge r_2)\Big\} + \\ + \Big\{\sigma_1 I(X_{j-1} \le r_1) + \sigma_2 I\Big(r_2 > X_{j-1} \ge r_1\Big) + \sigma_3 I(X_{j-1} > r_2)\Big\}(W_j - W_{j-1}), \end{gathered} \tag{3}$$

where W_j increment of the Wiener process, $W_j - W_{j-1} \sim N(0, \Delta_j)$ and the time step of discretization $\Delta_j = t_j - t_{j-1}$. If the values of X_0 and r_1, r_2 are available, then from expression (3) we can obtain the function of doubled negative log-likelihood $-2l_x(\theta)$.

As a result, for each value of X_j, given the previous value of X_{j-1}, the approximation provides a normal distribution of the increment

$$\begin{gathered}-2l_x(\theta) = C + \sum_{j=1}^{q} log\Big\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I\Big(r_2 > X_{j-1} \ge r_1\Big) + \big(\sigma_3^2 I(X_{j-1} > r_2)\ \Big\} + \\ + \sum_{j=1}^{q} \frac{\left[X_j - X_{j-1} - \Delta_j \left\{\begin{matrix}(\beta_{10} + \beta_{11}X_{j-1})I(X_{j-1} \le r_1) + (\beta_{20} + \beta_{21}X_{j-1})I(r_2 > X_{j-1} \ge r_1) + \\ (\beta_{30} + \beta_{31}X_{j-1})I(X_{j-1} > r_2)\end{matrix}\right\}\right]^2}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j},\end{gathered} \tag{4}$$

for some constant C, which does not affect the estimation by the maximum likelihood method.

The system of equations for estimating the model parameters will include the following model parameters: $\beta_{10}, \beta_{11}, \beta_{20}, \beta_{21}, \beta_{30}, \beta_{31}, \sigma_1^2, \sigma_2^2, \sigma_3^2$, for thresholds r_1 and r_2.

Differentiating the equation by the shift parameters:

$\boldsymbol{\beta_{10}, \beta_{11}}$ **at** $\boldsymbol{X_{j-1} \le r_1}$:

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{10}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{10} + \beta_{11}X_{j-1})\big)I(X_{j-1} \le r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j};$$

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{11}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{10} + \beta_{11}X_{j-1})\big)X_{j-1}I(X_{j-1} \le r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j}.$$

$\boldsymbol{\beta_{20}, \beta_{21}}$ **at** $\boldsymbol{r_2 > X_{j-1} \ge r_1}$:

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{20}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{20} + \beta_{21}X_{j-1})\big)I(r_2 > X_{j-1} \ge r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j};$$

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{21}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{20} + \beta_{21}X_{j-1})\big)X_{j-1}I(r_2 > X_{j-1} \ge r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j}.$$

$\boldsymbol{\beta_{30}, \beta_{31}}$ **at** $\boldsymbol{X_{j-1} > r_2}$:

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{30}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{30} + \beta_{31}X_{j-1})\big)I(X_{j-1} > r_2)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j};$$

$$-2\frac{\partial l_X(\theta)}{\partial \beta_{31}} = -2\sum_{j=1}^{q} \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{30} + \beta_{31}X_{j-1})\big)X_{j-1}I(X_{j-1} > r_2)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}\Delta_j}.$$

Differentiation of the equation with respect to the diffusion parameters σ_1^2, σ_2^2, σ_3^2:

$$\begin{aligned}-2\frac{\partial l_X(\theta)}{\partial \sigma_1^2} = \sum_{j=1}^{q} & \frac{I(X_{j-1} \le r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}} \\ & - \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{10} + \beta_{11}X_{j-1})\big)^2 I(X_{j-1} \le r_1)}{\{\sigma_1^4 I(X_{j-1} \le r_1) + \sigma_2^4 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^4 I(X_{j-1} > r_2)\ \}\Delta_j}; \\ -2\frac{\partial l_X(\theta)}{\partial \sigma_2^2} = \sum_{j=1}^{q} & \frac{I(r_2 > X_{j-1} \ge r_1)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \le r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}} \\ & - \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{20} + \beta_{21}X_{j-1})\big)^2 I(r_2 > X_{j-1} \ge r_1)}{\{\sigma_1^4 I(X_{j-1} \le r_1) + \sigma_2^4 I(r_2 > X_{j-1} \le r_1) + (\sigma_3^4 I(X_{j-1} > r_2)\ \}\Delta_j}; \\ -2\frac{\partial l_X(\theta)}{\partial \sigma_3^2} = \sum_{j=1}^{q} & \frac{I(X_{j-1} > r_2)}{\{\sigma_1^2 I(X_{j-1} \le r_1) + \sigma_2^2 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^2 I(X_{j-1} > r_2)\ \}} \\ & - \frac{\big(X_j - X_{j-1} - \Delta_j(\beta_{30} + \beta_{31}X_{j-1})\big)^2 I(X_{j-1} > r_2)}{\{\sigma_1^4 I(X_{j-1} \le r_1) + \sigma_2^4 I(r_2 > X_{j-1} \ge r_1) + (\sigma_3^4 I(X_{j-1} > r_2)\ \}\Delta_j}.\end{aligned}$$

By equating the results of differentiation by parameters to zero, we obtain a system of equations for the parameters in matrix form:

$$
\begin{aligned}
\begin{bmatrix}\beta_{10}\\ \beta_{11}\end{bmatrix} &= \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \\ \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-2}} \end{bmatrix}^{-1} \\
&* \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} \\ \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \end{bmatrix}; \\
\begin{bmatrix}\beta_{20}\\ \beta_{21}\end{bmatrix} &= \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{\Delta_j I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \\ \sum\limits_{i=1}^{q} \frac{\Delta_j I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-2}} \end{bmatrix}^{-1} \\
&* \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} \\ \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \end{bmatrix}; \\
\begin{bmatrix}\beta_{30}\\ \beta_{31}\end{bmatrix} &= \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \\ \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} & \sum\limits_{i=1}^{q} \frac{\Delta_j I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-2}} \end{bmatrix}^{-1} \\
&* \begin{bmatrix} \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}} \\ \sum\limits_{i=1}^{q} \frac{(X_j-X_{j-1})I(X_{j-1}>r_2)}{\{\sigma_1^2 I(X_{j-1}\le r_1)+\sigma_2^2 I(r_2>X_{j-1}\ge r_1)+\sigma_3^2 I(X_{j-1}>r_2)\}X_{j-1}^{-1}} \end{bmatrix}; \\
\sigma_1^2 &= \frac{\sum_{i=1}^{q} \frac{\{X_j-X_{j-1}-\Delta_j(\beta_{10}+\beta_{11}X_{j-1})\}^2 I(X_{j-1}\le r_1)}{\{\sigma_1^4 I(X_{j-1}\le r_1)+\sigma_2^4 I(r_2>X_{j-1}\ge r_1)+\sigma_3^4 I(X_{j-1}>r_2)\}\Delta_j}}{\sum\limits_{i=1}^{q} \frac{I(X_{j-1}\le r_1)}{\{\sigma_1^2 I(X_{j-1}\le r_1)\}}}; \\
\sigma_2^2 &= \frac{\sum_{i=1}^{q} \frac{\{X_j-X_{j-1}-\Delta_j(\beta_{20}+\beta_{21}X_{j-1})\}^2 I(r_2>X_{j-1}\ge r_1)}{\{\sigma_1^4 I(X_{j-1}\le r_1)+\sigma_2^4 I(r_2>X_{j-1}\ge r_1)+\sigma_3^4 I(X_{j-1}>r_2)\}\Delta_j}}{\sum\limits_{i=1}^{q} \frac{I(r_2>X_{j-1}\ge r_1)}{\{\sigma_2^2 I(r_2>X_{j-1}\ge r_1)\}}}; \\
\sigma_3^2 &= \frac{\sum_{i=1}^{q} \frac{\{X_j-X_{j-1}-\Delta_j(\beta_{30}+\beta_{31}X_{j-1})\}^2 I(X_{j-1}>r_2)}{\{\sigma_1^4 I(X_{j-1}\le r_1)+\sigma_2^4 I(r_2>X_{j-1}\ge r_1)+\sigma_3^4 I(X_{j-1}>r_2)\}\Delta_j}}{\sum\limits_{i=1}^{q} \frac{I(X_{j-1}>r_2)}{\{\sigma_3^2 I(X_{j-1}>r_2)\}}}.
\end{aligned}
\tag{5}
$$

Thus, the system of equations for parameter estimation consists of a separate selection of parameters for each mode using the given model. The results depend on the values of the thresholds r_1 and r_2, which divide the trajectory into subsets. These equations are used as the basis for an iterative algorithm for parameter and threshold estimation.

3. Iterative Algorithm for Estimating Parameters and Thresholds

The calculation of the parameters $\hat{\beta}_{10}, \hat{\beta}_{11}, \hat{\beta}_{20}, \hat{\beta}_{21}, \hat{\beta}_{30}, \hat{\beta}_{31}, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_3^2$ and the thresholds $\hat{r}_1$, $\hat{r}_2$ in the three-mode two-threshold Ornstein–Uhlenbeck process can be accomplished using the improved iterative algorithm (Figure 1) described in [8], as outlined below. $\hat{\beta}_{i0}, \hat{\sigma}_i^2, \hat{r}_i$ an estimated value of the parameters. The estimation procedure proceeds iteratively until convergence is achieved.

Step 1. Initialization.

Time series $\{X_0, X_1, \ldots, X_q\}$, for which the ordinal statistics $\left\{X_{(0)}, X_{(1)}, \ldots, X_{(q)}\right\}$ are calculated, where $X_{(0)} \le X_{(1)} \le \ldots \le X_{(q)}$

Step 2. Iterative Estimation.

2.1. It is necessary to fix the threshold vector $\hat{r}$ for $i = 0, 1, \ldots, \lambda$, and initialize the variance estimates $\hat{\sigma}_1^2(0), \hat{\sigma}_2^2(0), \hat{\sigma}_3^2(0)$. λ is the index limit (the number of tested parameter combinations).

2.2. For each iteration $k \geq 1$, we update the parameter vector

$$\hat{\theta}(k) = \left(\hat{\beta}_{10}(k), \hat{\beta}_{11}(k), \hat{\beta}_{20}(k), \hat{\beta}_{21}(k), \hat{\beta}_{30}(k), \hat{\beta}_{31}(k), \hat{\sigma}_1^2(k), \hat{\sigma}_2^2(k), \hat{\sigma}_3^2(k)\right)$$

recursively using the update equations (see Equation (5)). In each equation, the previous iteration's values $\hat{\beta}_{ij}$ by $\hat{\beta}_{ij}(k-1)$, $\hat{\sigma}_i^2$ by $\hat{\sigma}_i^2(k-1)$, i = 1, 2, 3; j = 0, 1.

2.3. Repeat Step 2.2 until the estimated value of the parameter vector $\hat{\theta}(k)$ is convergent. In our implementation, convergence is assessed using a tolerance of $\varepsilon = 10^{-6}$ and the spectral analysis of the Jacobian matrix (see Section 4 for details).

Let the convergent estimate be of the form

$$\hat{\theta}_i = \left(\hat{\beta}_{10,i}(k), \hat{\beta}_{11,i}(k), \hat{\beta}_{20,i}(k), \hat{\beta}_{21,i}(k), \hat{\beta}_{30,i}(k), \hat{\beta}_{31,i}(k), \hat{\sigma}_{1,i}^2(k), \hat{\sigma}_{2,i}^2(k), \hat{\sigma}_{3,i}^2(k)\right).$$

Step 3. Model Selection.

For each $i = 0, 1, \ldots, \lambda$, compute the negative log-likelihood value $-2l_X(\hat{\theta}_i)$. Then select the index that gives the highest likelihood $\tau = \text{argmin}_i\{-2l_X(\hat{\theta}_i)\}$ and define the approximate maximum likelihood estimate of the full parameter set $\eta = (r, \theta)$ as $\hat{\eta} = (\hat{r}_\tau, \hat{\theta}_\tau)$.

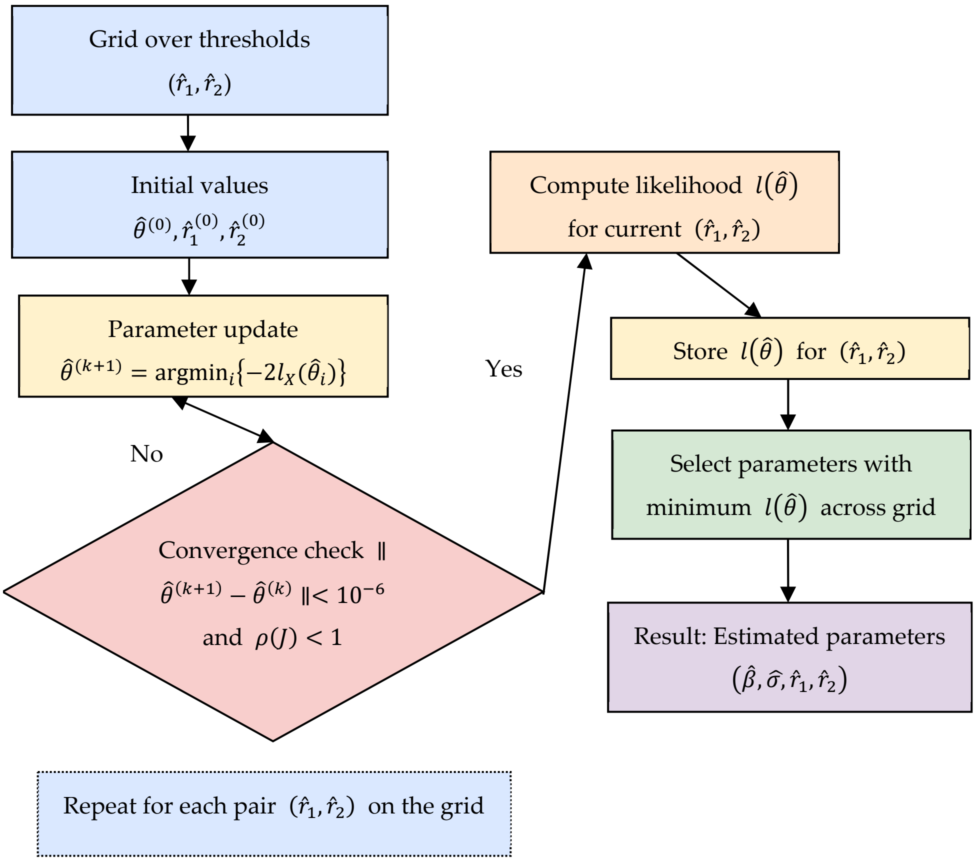

Figure 1. Block diagram of the Iterative Algorithm for Estimating Parameters and Thresholds.

4. Iterative Parameter Estimation Procedure

The parameter estimation for the two-threshold Ornstein–Uhlenbeck (OU) process with three dynamic regimes involves solving a complex optimization problem due to the presence of non-linearities and regime-dependent structural breaks. The model assumes that the observed stochastic process X_t transitions between three distinct modes, each governed by its own drift and diffusion coefficients, with the transitions determined by two unknown thresholds $\hat{r}_1$ and $\hat{r}_2$. The parameter vector to be estimated is defined as

$$\theta = \begin{pmatrix} \hat{\beta}_{10} \\ \hat{\beta}_{11} \\ \hat{\beta}_{20} \\ \hat{\beta}_{21} \\ \hat{\beta}_{30} \\ \hat{\beta}_{31} \\ \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{\sigma}_3^2 \end{pmatrix},$$

where $\hat{\beta}_{ij}$ are the drift parameters and $\hat{\sigma}_1^2$ are the regime-specific diffusion variances for i = 1, 2, 3 and j = 0, 1. The model captures the piecewise linear behavior of the system dynamics over the intervals $(-\infty, \hat{r}_1]$, $(\hat{r}_1, \hat{r}_2]$, and $(\hat{r}_1, \infty)$. The initial thresholds r_1, r_2 are generated over a uniform grid within the range of observed data X_t, excluding 5% at both ends to avoid empty regimes. For each pair r_1, r_2, the initial variance parameters σ_i^2 are set to half of the empirical variance of observations within the corresponding region. The regression coefficients β_i are then estimated analytically in the first iteration using the normal equations.

This data-driven initialization provides stable starting values and ensures convergence without requiring arbitrary parameter assumptions.

To estimate these parameters, an iterative numerical procedure is employed, leveraging the structure of the Approximate Maximum Likelihood Estimation (AMLE). At the core of the algorithm lies a fixed-point iteration scheme, where the current estimate of the parameter vector $\hat{\theta}^{(k)}$ at iteration k is updated via a transformation function F, such that

$$\hat{\theta}^{(k+1)} = F\left(\hat{\theta}^{(k)}\right).$$

The mapping F is defined implicitly by the likelihood function, and incorporates regime segmentation, parameter re-estimation, and variance update steps. Specifically, each iteration proceeds as follows:

- ***Segmentation.*** The dataset is partitioned into three subsets based on current estimates of the thresholds $\hat{r}_1^{(k)}$ and $\hat{r}_2^{(k)}$. This segmentation defines the data points corresponding to each regime.
- ***Parameter Update.*** Within each regime, drift parameters $\hat{\beta}_{ij}$ are estimated via linear regression using least-squares or generalized method of moments, and diffusion parameters $\hat{\sigma}_i^2$ are updated based on the empirical variance of the residuals. These estimators exploit the local linearity of the OU process and assume conditional independence of the noise terms across time steps.
- ***Threshold Optimization.*** Updated thresholds may be obtained via grid search or gradient-based optimization techniques aimed at maximizing the approximate likelihood. Alternatively, a profile likelihood approach may be employed where the likelihood is maximized over thresholds for each fixed $\hat{\theta}^{(k)}$.
- ***Convergence Check.*** The process continues until the difference between consecutive parameter estimates satisfies a predefined convergence criterion:

$$\| \hat{\theta}^{(k+1)} - \hat{\theta}^{(k)} \| < \varepsilon.$$

Here, ε predefined the convergence tolerance, or until the log-likelihood stabilizes:

$$\left| l_X\left(\hat{\theta}^{(k+1)}\right) - l_X\left(\hat{\theta}^{(k)}\right) \right| < \delta.$$

Here, δ predefined the threshold for the change in the log-likelihood. The function F is constructed to ensure convergence under suitable regularity conditions. In particular, if F is a contraction mapping in the relevant parameter space, then by the *Banach Fixed Point Theorem*, the sequence $\left\{\hat{\theta}^{(k)}\right\}$ will converge to a unique fixed point θ^*, which corresponds to the optimal parameter estimates.

To confirm this, the *Jacobian matrix* of partial derivatives,

$$J_{ij} = \frac{\partial F_i\left(\hat{\theta}\right)}{\partial \hat{\theta}_j},$$

is approximated numerically using a forward finite-difference scheme. J_{ij} denotes the element of the Jacobian matrix corresponding to the partial derivative of the i-th function $F_i\left(\hat{\theta}\right)$ with respect to the j-th parameter $\hat{\theta}_j$. The *spectral radius* $\rho(J)$ of the Jacobian, defined as the largest modulus of its eigenvalues, provides a sufficient condition for convergence [23]:

$$\rho(J) < 1 \Longrightarrow \text{contractive mapping} \Longrightarrow \text{guaranteed convergence.}$$

This iterative framework allows for a flexible and computationally efficient approach to parameter estimation in threshold-based diffusion processes. The iterative procedure is terminated when both numerical and spectral convergence criteria are satisfied. Specifically, the step size between successive parameter estimates is controlled by a fixed tolerance $\varepsilon = 10^{-6}$ [24]. By dynamically reassigning regimes and updating parameters, the algorithm adapts to the inherent non-linearity and discontinuity in the observed time series data. The procedure is particularly suitable for applications where regime transitions are unexpected and the dynamics within each regime are sufficiently different.

To prove that the calculations are close to finding the correct result, we will use the fixed point method: if the function F at each subsequent iteration reduces the difference between the values, then in the future everything will converge to a stable result, i.e.,

$$\left\| F\left(\hat{\theta}^{(1)}\right) - F\left(\hat{\theta}^{(2)}\right) \right\| \le q \left\| \hat{\theta}^{(1)} - \hat{\theta}^{(2)} \right\|, \;\; 0 < q < 1,$$

where q is the contraction coefficient. Then, there is a single fixed point $\boldsymbol{\theta}^*$, and the sequence $\boldsymbol{\theta}^{(k)}$ will gradually approach it. To verify this property, we can use the spectral analysis of the Jacobi matrix, i.e., each element shows how much the parameter $\boldsymbol{\theta_i}$ changes when the parameter $\boldsymbol{\theta_j}$ changes.

For a model with two thresholds, the Jacobi matrix has a size of 9×9, for example

$$J = \begin{pmatrix} \frac{\partial \hat{\beta}_{10}^{(k+1)}}{\partial \hat{\beta}_{10}^{(k)}} & \cdots & \frac{\partial \hat{\beta}_{10}^{(k+1)}}{\partial \hat{\sigma}_3^{2,\,(k)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{\sigma}_3^{2,(k+1)}}{\partial \hat{\beta}_{10}^{(k)}} & \cdots & \frac{\partial \hat{\sigma}_3^{2,(k+1)}}{\partial \hat{\sigma}_3^{2,\,(k)}} \end{pmatrix}.$$

To estimate the Jacobi matrix at each step of the iterative algorithm, instead of analytical derivatives, a numerical approximation method based on a one-way finite-difference scheme (forward-difference [25,26]) was used:

$$J_{ij} \approx \frac{\hat{\theta}_i^{(k+1)} - \hat{\theta}_i^{(k)}}{\hat{\theta}_j^{(k)}} \left(\forall\, \hat{\theta}_j^{(k)} \neq 0\right).$$

The numerator is the change in the value of $\boldsymbol{\theta_i}$ between iterations, and the denominator is the previous value of the parameter $\boldsymbol{\theta_j}$, i.e., an approximation of the relative change in one parameter relative to another. We restrict the domain of J_{ij} to $\left|\hat{\theta}_j^{(k)}\right| \geq \varepsilon$ (with a small floor ε) this yields a stable local Jacobian proxy used for the spectral-radius check $\rho(J) < 1$.

After constructing the approximate Jacobi matrix J, its spectral radius is calculated—the maximum modulus eigenvalue among all eigenvalues of the matrix $\boldsymbol{\rho(J) = max|\lambda_i|}$ If $\boldsymbol{\rho(J) < 1}$, then the iterative process converges to a stable solution (the Banach theorem confirms this). If $\boldsymbol{\rho(J)} > 1$, then convergence is not guaranteed.

To substantiate the convergence, numerical analysis was used, which makes it possible to check whether with each step of the iterative algorithm the parameter update function makes their values closer to each other. The described approach allows mathematically justifying the performance of the proposed model. We clarify that the parameter sequence $\left\{\hat{\theta}(k)\right\}$ is generated by the map $\hat{\theta}^{(k+1)} = G\left(\hat{\theta}^{(k)}; \hat{r}_1, \hat{r}_2\right)$. The function G represents an iterative update rule for estimating the threshold process parameters (it describes one iterative step of the algorithm). Under standard assumptions (continuity and differentiability of G, local identifiability), if the spectral radius of the Jacobian at a fixed point satisfies $p(J_G(\theta^*)) < 1$, then G is a contraction in a neighborhood of θ^*, and $\hat{\theta}^{(k)} \to \theta^*$ (Banach fixed-point theorem). Our implementation explicitly verifies $p\left(J^{(k)}\right) < 1$ and a step-size tolerance $\| \hat{\theta}^{(k+1)} - \hat{\theta}^{(k)} \| < 10^{-6}$ to declare convergence. Given the non-convex likelihood, we claim local (not global) convergence and employ multi-start initializations to reduce sensitivity to starting values.

The convergence of the proposed iterative algorithm for estimating the parameters of the three-mode two-threshold Ornstein–Uhlenbeck process may include several factors: the initial initialization of the threshold values and the algorithm parameters (stopping criterion). The correct convergence assessment ultimately affects the number of iterations and the accuracy of the parameter calculation.

Initial threshold values. Incorrect selection of the initial threshold values can lead to slow operation of the algorithm, since many iterations will need to be performed. Threshold values close to the expected optimal values usually accelerate convergence. The algorithm demonstrates sensitivity to the variance level of the process: with small σ (low volatility), trajectories become smoother and provide less information for identifying threshold regimes, while large σ (high volatility) increases variability in estimates due to frequent regime switching and stronger stochastic fluctuations.

Stopping criterion. From the set conditions for completing the iterative process. If the convergence criterion is set too large, the algorithm may terminate and as a result the accuracy of such calculations will be at a low level. If the convergence criterion is set too small, more iterations will be required for calculations, and it will work even with minor changes in the parameters, which will lead to the use of extra time for calculations. The optimal convergence criterion should provide a balance between the accuracy of the result and the time and resources for its calculation. Several criteria can be defined to assess the convergence between iterations:

- Estimation of the change in parameters between iterations. The parameter values are compared in the current and previous iterations. A slight difference in the values of the calculated parameters indicates that the algorithm has approached the optimal solution, for example, $\hat{\sigma}_i$ for iteration $(k+1)$ minus $\hat{\sigma}_i$ for iteration (k).

- Change in the estimate of the likelihood function between iterations $\left|l_X\left(\hat{\theta}_{k+1}\right) - l_X\left(\hat{\theta}_k\right)\right|$. If the difference between the values approaches zero, then the following steps do not improve the model and the threshold value is fixed at which the maximum likelihood is achieved and further optimization is impractical.

5. Numerical Implementation and Results

An R-based computational implementation was developed to execute the proposed iterative estimation algorithm for the two-threshold Ornstein–Uhlenbeck process. We introduce the pseudocode of Algorithm 1.

Algorithm 1. Two-threshold Ornstein–Uhlenbeck Process Estimation

```
SortedX = Sort(X)
for I = 0 to Lambda do
OrderStatistic[I] = SortedX[I]
end for
    for I = 0 to Lambda do
        R[I] = FixValue(I)
        InitialParameters[I] = InitializeModelParameters()
end for
    for I = 0 to Lambda do
        K = 1
        Converged = false
        while not Converged and K ≤ N_Max do
                Theta[I][K] = EstimateTheta(
                Theta[I][K − 1], R[I], Epsilon
                )
    if Norm(Theta[I][K] − Theta[I][K − 1]) < Epsilon then
        Converged = true
    else
                K = K + 1
    end if
  end while
 end for
 for I = 0 to Lambda do
    LogLikelihood[I] = ComputeNeg2LogLikelihood(X, Theta[I][K])
 end for
    Tau = ArgMin(LogLikelihood)
  Ny = (R[Tau], Theta[Tau])
```

To assess the performance and convergence behavior of the method, a synthetic time series of length $n = 100$ was generated, simulating a stochastic process with three distinct dynamic regimes. The constructed series included: a decreasing trend in the first segment ($n_1 = 30$ observations), a near-stationary behavior in the second segment ($n_2 = 30$), and an increasing trend in the final segment ($n_3 = 40$). Each regime was characterized by a linear deterministic drift component and superimposed Gaussian noise, where the variance of the noise was varied between regimes to test the sensitivity of the model to heteroscedasticity. These regimes were concatenated to form the complete dataset $\{X_t\}_{t=1}^{n}$, intended to represent a process transitioning across structural thresholds with distinct statistical properties.

Code written in R-4.3.0 using both standard packages and hand-built procedures and functions implemented the iterative estimation routine, which included dynamic segmentation based on candidate thresholds, parameter updates via approximate maximum likelihood equations, and convergence checks. The thresholds $\hat{r}_1$ and $\hat{r}_2$ were initialized using empirical quantiles of the time series and subsequently optimized as part of the iterative procedure.

During execution, the algorithm successfully recovered the underlying drift coefficients $\hat{\beta}_{ij}$ and diffusion variances $\hat{\sigma}_i^2$ corresponding to each regime. The estimated thresholds $\hat{r}_1$ and $\hat{r}_2$ closely matched the true breakpoints embedded in the synthetic data generation process, thereby validating the method's ability to detect transition points between different dynamical behaviors.

Convergence of the parameter estimates was evaluated using two criteria: the stability of parameter values across successive iterations (based on a predefined tolerance) and the monotonic behavior of the log-likelihood function. In all test cases, the iterative algorithm demonstrated stable convergence. Specifically, there was a negligible change in both the parameter vector $\hat{\theta}^{(k)}$ and the likelihood value after a finite number of iterations. These observations confirm that the fixed-point mapping defined by the algorithm is contractive within the neighborhood of the true parameter values. Therefore, it satisfies the necessary conditions for convergence.

Additionally, graphical visualization of the time series and the estimated thresholds was used to verify the accuracy of regime segmentation. The plots revealed a clear correspondence between the detected change points and the synthetic transitions in drift behavior, which confirms the robustness of the estimation framework. This simulation-based validation provides a strong basis for future application of the algorithm to real-world data in which structural thresholds are unknown.

To test the algorithm at the software level, a time series was generated that describes the behavior of the two-mode Ornstein–Uhlenbeck pre-threshold process with different dynamics in three intervals n_1, n_2 and n_3.

Figures 2–6 display the time series trajectories X_t, and displays the optimal value of the thresholds that record the transition between modes. The generated time series included a linear shift (characteristic for each data slice) and was supplemented with Gaussian noise with different variance. The entire dataset was combined into a set of values X_t.

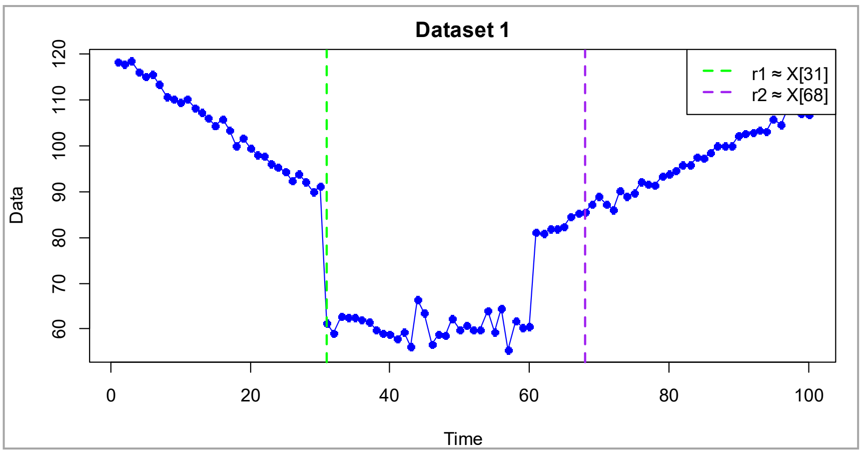

Figure 2. Synthetic dataset 1 (dimensionless). Blue dots represent simulated observations. Vertical dashed lines show thresholds $r_1 \approx X[31]$ and $r_2 \approx X[68]$, dividing the process into three regimes: n_1 (decreasing, $X_t < r_1$), n_2 (nearly stationary, $r_1 \leq X_t \leq r_2$), and n_3 (increasing, $X_t > r_2$).

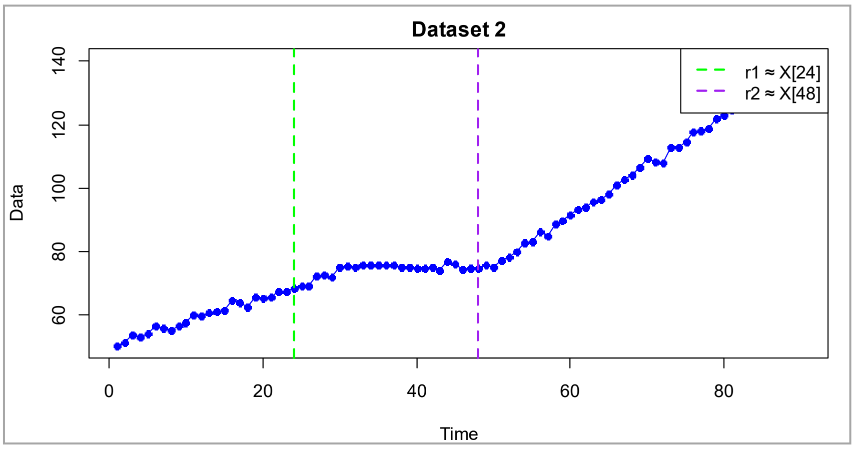

Figure 3. Synthetic dataset 2 (dimensionless). Blue dots represent simulated observations. Vertical dashed lines show thresholds $r_1 \approx X[24]$ and $r_2 \approx X[48]$, dividing the process into three regimes: n_1 (increasing, $X_t < r_1$), n_2 (nearly stationary, $r_1 \leq X_t \leq r_2$), and n_3 (increasing, $X_t > r_2$).

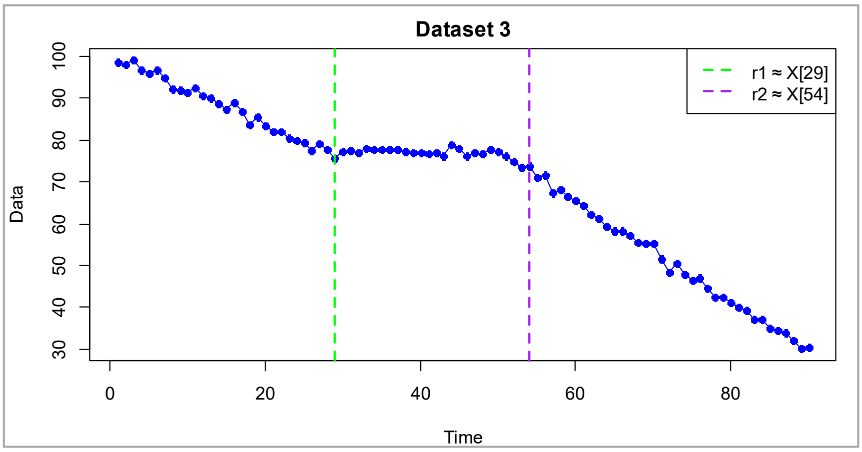

Figure 4. Synthetic dataset 3 (dimensionless). Blue dots represent simulated observations. Vertical dashed lines show thresholds $r_1 \approx X[29]$ and $r_2 \approx X[54]$, dividing the process into three regimes: n_1 (deceasing, $X_t < r_1$), n_2 (nearly stationary, $r_1 \leq X_t \leq r_2$), and n_3 (decreasing, $X_t > r_2$).

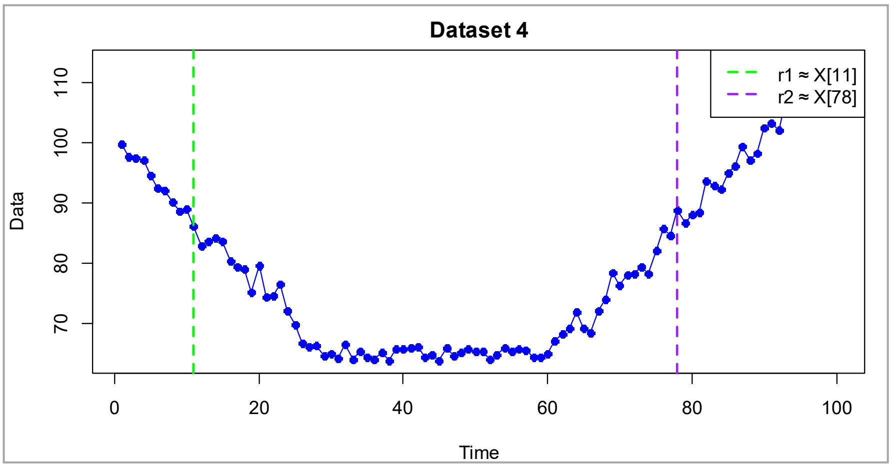

Figure 5. Synthetic dataset 4 (dimensionless). Blue dots represent simulated observations. Vertical dashed lines show thresholds $r_1 \approx X[11]$ and $r_2 \approx X[78]$, dividing the process into three regimes: n_1 (decreasing, $X_t < r_1$), n_2 (nearly stationary, $r_1 \leq X_t \leq r_2$), and n_3 (increasing, $X_t > r_2$).

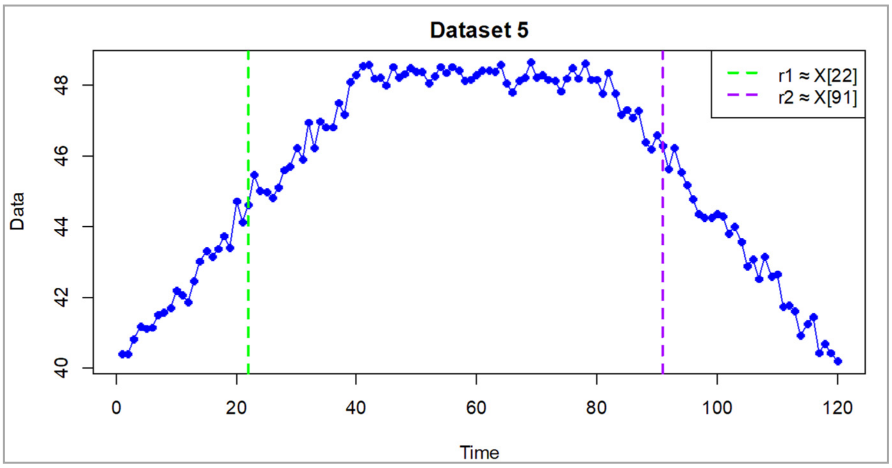

Figure 6. Synthetic dataset 5 (dimensionless). Blue dots represent simulated observations. Vertical dashed lines show thresholds $r_1 \approx X[22]$ and $r_2 \approx X[91]$, dividing the process into three regimes: n_1 (increasing, $X_t < r_1$), n_2 (nearly stationary, $r_1 \leq X_t \leq r_2$), and n_3 (decreasing, $X_t > r_2$).

Comparison between approximate true (True, approx.) and estimated (Estimated) threshold values $\hat{r}_1$ and $\hat{r}_2$ for five synthetic datasets Table 1. The true thresholds were determined from the transition points between the three regimes (n_1, n_2, n_3) in the generated time data. In most cases, the estimated parameters are close to the original ones, confirming the stability of the estimation algorithm; minor discrepancies can be attributed to the stochastic nature of the process and the non-convex shape of the likelihood surface.

Table 1. Approximate and estimated threshold.

Dataset	Threshold	True (Approx.)	Estimated	Abs. Error	Rel. Error (%)
Dataset 1	r_1	61.27939	61.27939	0.00000	0.00
Dataset 1	r_2	81.07964	85.65300	4.57336	5.64
Dataset 2	r_1	75.25381	68.47111	6.78270	9.01
Dataset 2	r_2	75.18712	74.88049	0.30663	0.41
Dataset 3	r_1	77.25381	75.66186	1.59195	2.06
Dataset 3	r_2	77.18712	73.82242	3.36470	4.36
Dataset 4	r_1	80.50946	86.34336	5.83390	7.25
Dataset 4	r_2	82.24510	89.07161	6.8265	8.30
Dataset 5	r_1	48.29144	44.62014	3.67130	7.60
Dataset 5	r_2	48.15795	46.27934	1.87861	3.90

The obtained results indicate that the algorithm reconstructs the threshold values with a small absolute error (within 0–6.8) and a relative error not exceeding 9%, confirming the high accuracy of parameter estimation across different datasets.

The graphs below, based on Dataset 4, illustrate the results from a different analytical perspective, providing a deeper view of the process dynamics and its structural features.

Figure 5 shows the time series of the observed process together with the estimated thresholds $\hat{r}_1$ and $\hat{r}_2$. The green dashed line corresponds to the lower threshold $\hat{r}_1 \approx 86.34$, which is localized within the descending segment of the trajectory, while the purple line represents the upper threshold $\hat{r}_2 \approx 89.07$, associated with the phase where the process begins to rise. This threshold configuration intuitively aligns with the phase behavior of the series: the lower threshold separates the declining regime, whereas the upper one marks the recovery phase and the transition to growth.

Figure 7 presents the likelihood function map in the space of the threshold parameters $\hat{r}_1$ and $\hat{r}_2$. The color scale represents the normalized log-likelihood value averaged per observation: lighter areas correspond to higher (worse) criterion values, while darker regions indicate a better model fit to the data. The red dot marks the minimum corresponding to the estimated optimal thresholds. As seen from the figure, the surface exhibits a distinctly non-convex structure: most of the parameter space yields a poor fit, whereas the global minimum is concentrated within a narrow, localized region.

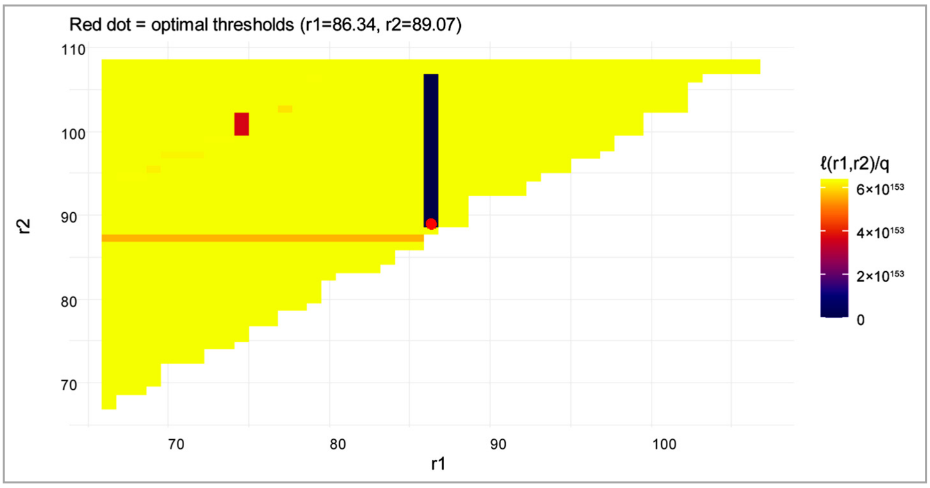

Figure 7. Average log-likelihood per observation (normalized).

Thus, the combination of these two illustrations demonstrates, first, the interpretability of the estimated thresholds in the time-series dynamics and, second, confirms the validity of the chosen optimization procedure even under a non-convex likelihood surface. This highlights the effectiveness of the proposed approximate maximum likelihood method for estimating the parameters of the two-threshold model.

Figure 8 illustrates the three-dimensional landscape of the likelihood function in the parameter space of $\hat{r}_1$ and $\hat{r}_2$. The red dot indicates the optimal threshold values $\hat{r}_1$ and $\hat{r}_2$ corresponding to the global minimum of the function. This representation clearly demonstrates the shape of the surface, the location of local and global minima, and allows for a visual assessment of the structure (non-convexity) of the optimization problem.

Figure 9 shows cross-sections of the likelihood surface for fixed values of $\hat{r}_1$. As $\hat{r}_2$ varies, the function exhibits several local minima, confirming its non-convex nature. This explains why the estimation results may be sensitive to the initial parameter values and require careful selection of starting points or multiple runs of the algorithm with different initializations.

Figure 10 shows how the number of iterations required for convergence depends on the average spectral radius of the Jacobian matrix $\rho(J)$. Blue dots represent the results for different pairs of thresholds $(\hat{r}_1,\ \hat{r}_2)$, while the red curve shows a smoothed approximation illustrating the general trend of the relationship between the convergence rate and the spectral radius. It can be seen that for small values $p(J) < 0.4$, the algorithm typically converges within 3–4 iterations, whereas as the spectral radius increases, the number of iterations rises significantly. This confirms the theoretical statement that the convergence rate of the iterative process is inversely related to the spectral radius of the Jacobian matrix.

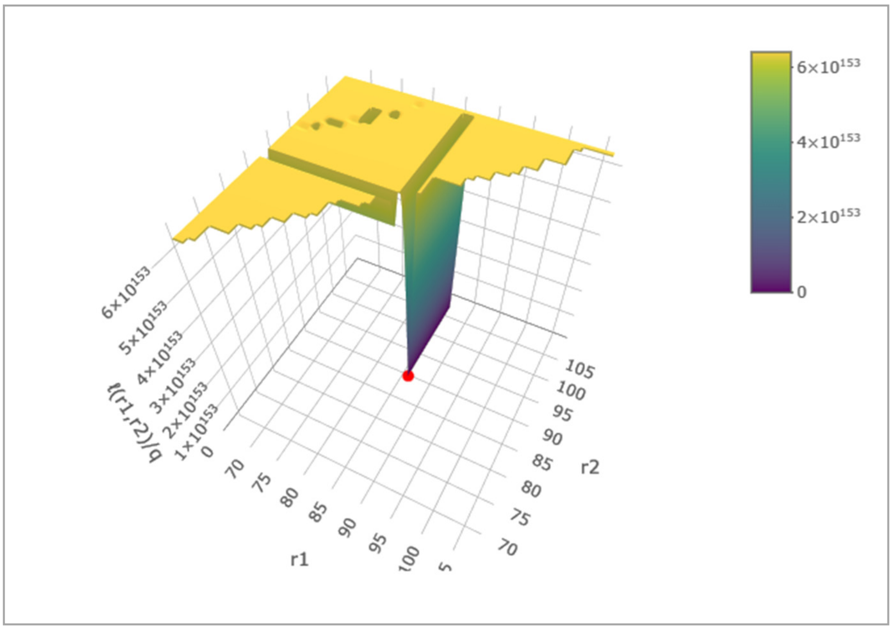

Figure 8. Three-dimensional likelihood surface with optimal thresholds.

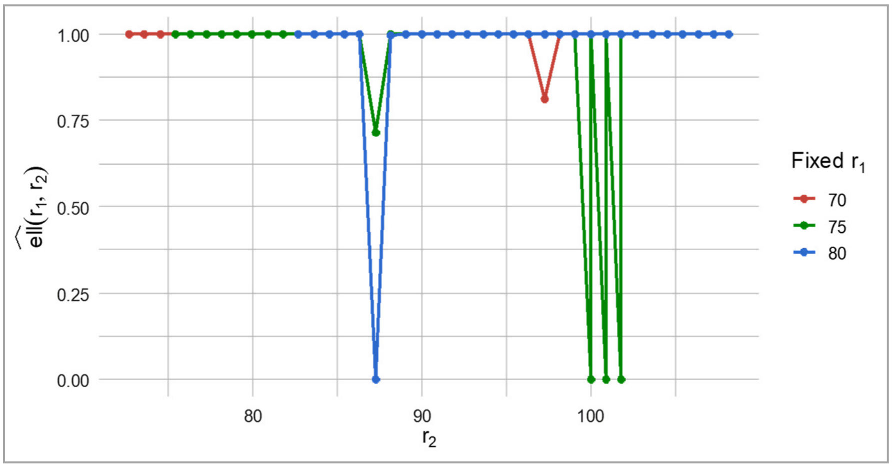

Figure 9. Cross-sections of the likelihood surface at fixed $\hat{r}_1$ values.

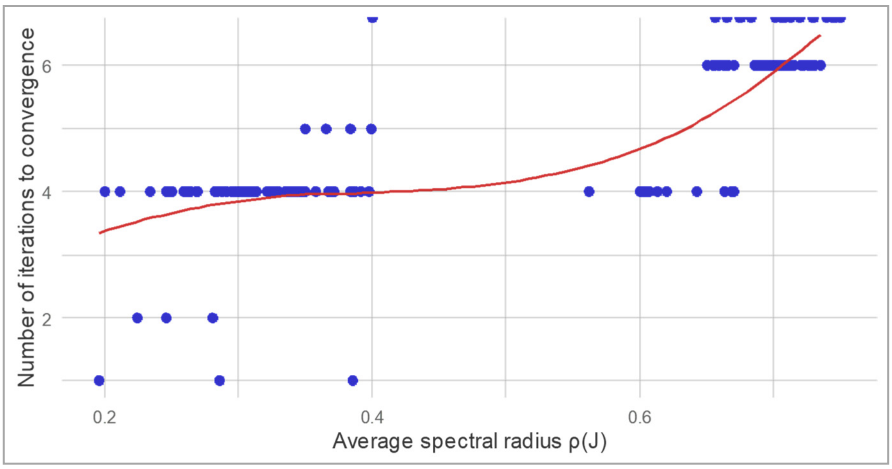

Figure 10. Dependence of the convergence rate on the spectral radius.

Using dataset 4 as an example, we assess several quantitative indicators of the model's accuracy in reproducing the process dynamics, namely the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Normalized Root Mean Squared Error (NRMSE).

According to the simulation results for dataset 4, the root mean squared error of increments was RMSE = 2.04, and the mean absolute error was MAE = 1.56, which is comparable to the standard deviation of noise $\sigma \approx 1.5$. Thus, the method adequately reproduces the stochastic component of the data.

The normalized error (NRMSE $\approx$ 0.97) indicates a close correspondence between the model and the empirical increments. The lowest accuracy is observed in the middle (second) regime, where the number of observations is minimal ($n_2 = 6$).

The algorithm achieved convergence after 8 iterations for the set of parameters corresponding to the minimum value of the likelihood function. The computation time directly depends on the processing power of the personal computer on which the program code is executed.

These visualizations confirm the model's ability to accurately detect regime changes, particularly when the signal-to-noise ratio is favorable. Higher dispersion may obscure threshold detection, but when transitions are distinct and noise is moderate, the algorithm reliably identifies structural shifts. Even under smooth trends and low noise, the method successfully captures changes in the data's dynamic structure, demonstrating flexibility and robustness across a range of scenarios.

6. Conclusions

This study demonstrated the effectiveness of using approximate maximum likelihood to calculate the parameters of a two-mode pre-threshold process. The developed methodology made it possible not only to calculate the shift and dispersion parameters for the Ornstein–Uhlenbeck process modes, but also to determine the threshold points r_1 and r_2, which divided the process dynamics into different segments with different properties. The implementation of the iterative algorithm in the program code confirmed the correctness of the calculations, and also revealed sensitivity to changes in the process data structure.

Compared to the methods proposed in [8], where a single-threshold model is considered, the pre-threshold model considered above allowed us to describe more complex, three-phase process dynamics. The study [9], which is based on the approximation of Lévy processes, illustrated the potential of using a two-threshold model in more complex stochastic processes.

As a result of this study, a logarithmic likelihood function with two thresholds was developed, the shift and dispersion parameters were estimated in the modes $X_t < r_1$, $r_1 \leq X_t < r_2$, $X_t \geq r_2$, an iterative algorithm for parameter estimation was implemented and its software implementation was developed.

Overall, the results confirmed that the proposed AMLE-based algorithm provided stable and accurate parameter recovery across different scenarios and noise levels. The findings provide a foundation for applying the two-threshold model to real-world time series data. This is particularly relevant in finance and related disciplines. The results also open opportunities for extending the approach to multidimensional threshold models based on the Ornstein–Uhlenbeck process.

Author Contributions: Conceptualization, S.B., A.N. and S.N.; methodology, A.N., S.B. and S.N.; formal analysis, S.B. and A.N.; investigation, A.N., S.B. and S.N.; writing—original draft preparation, A.N., S.B. and S.N.; writing—review and editing, A.N., S.B. and S.N.; project administration, S.B.; funding acquisition, S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of National Defence of Lithuania, Study Support Projects NoV-58, 30 January 2025, General Jonas ŽemaitisMilitary Academy of Lithuania, Vilnius, Lithuania.

Data Availability Statement: The complete dataset supporting the reported results is available at: https://zenodo.org/records/16282478.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chiou, Y.Y.; Su, L.; Tsai, H. Nonparametric Regression with Multiple Thresholds: Estimation and Inference. *J. Econom.* **2018**, *206*, 472–514. [CrossRef]
2. Scarffe, A.; Coates, A.; Brand, K.; Bikker, M.; Cowan, K. Decision Threshold Models in Medical Decision Making: A Scoping Literature Review. *BMC Med. Inform. Decis. Mak.* **2024**, *24*, 273. [CrossRef] [PubMed]
3. Lejay, A.; Pigato, P. A Threshold Model for Local Volatility: Evidence of Leverage and Mean Reversion Effects on Historical Data. *arXiv* **2017**, arXiv:1712.08329.
4. Rowland, M.; Tozer, T.N. *Clinical Pharmacokinetics and Pharmacodynamics: Concepts and Applications*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2011.
5. Gonzalo, J.; Pitarakis, J.-Y. Estimation and Model Selection Based Inference in Single and Multiple Threshold Models. *Rev. Econ. Stud.* **2002**, *69*, 813–833. [CrossRef]
6. Tong, H.; Lim, K.S. Threshold Autoregression, Limit Cycles and Cyclical Data. *J. R. Stat. Soc. Ser. B Methodol.* **1980**, *42*, 245–292. [CrossRef]
7. Uhlenbeck, G.E.; Ornstein, L.S. On the Theory of Brownian Motion. *Phys. Rev.* **1930**, *36*, 823–841. [CrossRef]
8. Yu, T.-H.; Tsai, H.; Rachinger, H. Approximate Maximum Likelihood Estimation of a Threshold Diffusion Process. *Comput. Stat. Data Anal.* **2020**, *142*, 106825. [CrossRef]
9. Tsai, H.; Nikitin, A.V. Threshold Models and Approximate Maximum Likelihood Estimation of Lévy Processes. *Cybern. Syst. Anal.* **2024**, *60*, 123–140. [CrossRef]
10. Milstein, G.N. *Numerical Integration of Stochastic Differential Equations*; Kluwer Academic Publishers: Boston, MA, USA, 1995.
11. Ji, L.; Li, C.; Zhou, X. Threshold Diffusions. *arXiv* **2025**, arXiv:2508.17812. [CrossRef]
12. Rachinger, H.; Lin, E.M.H.; Tsai, H. A Bootstrap Test for Threshold Effects in a Diffusion Process. *Comput. Stat.* **2023**, *38*, 901–919. [CrossRef]
13. Aït-Sahalia, Y. Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica* **2002**, *70*, 223–262. [CrossRef]
14. Chan, K.S. Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. *Ann. Stat.* **1993**, *21*, 520–533. [CrossRef]
15. Su, F.; Chan, K.S. Quasi-Likelihood Estimation of a Threshold Diffusion Process. *J. Econom.* **2015**, *189*, 473–484. [CrossRef]
16. Li, C. Maximum-Likelihood Estimation for Diffusion Processes via Closed-Form Density Expansions. *Ann. Stat.* **2013**, *41*, 1350–1380. [CrossRef]
17. Mazzonetto, S.; Pigato, P. Drift Estimation of the Threshold Ornstein–Uhlenbeck Process from Continuous and Discrete Observations. *arXiv* **2022**, arXiv:2008.12653v3. [CrossRef]
18. Tong, H. *Non-Linear Time Series: A Dynamical System Approach*; Oxford University Press: Oxford, UK, 1990.
19. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2013.
20. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the *EM* Algorithm. *J. R. Stat. Soc.* **1977**, *39*, 1–22. [CrossRef]
21. Skorohod, A.V. *Studies in the Theory of Random Processes*; Reprinted Edition; Dover Publications: New York, NY, USA, 1962.
22. Livina, V.N.; Lenton, T.M. A Modified Method for Detecting Incipient Bifurcations in a Dynamical System. *Geophys. Res. Lett.* **2007**, *34*, L03712. [CrossRef]
23. Kelley, C.T. *Iterative Methods for Linear and Nonlinear Equations*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1995; p. 7.
24. Burden, R.L.; Faires, J.D. *Numerical Analysis*, 9th ed.; Cengage Learning: Boston, MA, USA, 2011; pp. 60–62.

25. Kloeden, P.E.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 287–289.
26. Higham, D.J. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Rev.* **2001**, *43*, 525–546. [CrossRef]