

Article

Fractional Dynamics Identification via Intelligent Unpacking of the Sample Autocovariance Function by Neural Networks

Dawid Szarek ¹, Grzegorz Sikora ¹, Michał Balcerek ¹, Ireneusz Jabłoński ²
and Agnieszka Wyłomańska ^{1,*}

- ¹ Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland; 234856@student.pwr.edu.pl (D.S.); grzegorz.sikora@pwr.edu.pl (G.S.); michal.balcerek@pwr.edu.pl (M.B.)
- ² Department of Electronics, Wrocław University of Science and Technology, B. Prusa 53/55, 50-317 Wrocław, Poland; ireneusz.jablonski@pwr.edu.pl
- * Correspondence: agnieszka.wylomanska@pwr.edu.pl

Received: 31 October 2020; Accepted: 18 November 2020; Published: 20 November 2020



Abstract: Many single-particle tracking data related to the motion in crowded environments exhibit anomalous diffusion behavior. This phenomenon can be described by different theoretical models. In this paper, fractional Brownian motion (FBM) was examined as the exemplary Gaussian process with fractional dynamics. The autocovariance function (ACVF) is a function that determines completely the Gaussian process. In the case of experimental data with anomalous dynamics, the main problem is first to recognize the type of anomaly and then to reconstruct properly the physical rules governing such a phenomenon. The challenge is to identify the process from short trajectory inputs. Various approaches to address this problem can be found in the literature, e.g., theoretical properties of the sample ACVF for a given process. This method is effective; however, it does not utilize all of the information contained in the sample ACVF for a given trajectory, i.e., only values of statistics for selected lags are used for identification. An evolution of this approach is proposed in this paper, where the process is determined based on the knowledge extracted from the ACVF. The designed method is intuitive and it uses information directly available in a new fashion. Moreover, the knowledge retrieval from the sample ACVF vector is enhanced with a learning-based scheme operating on the most informative subset of available lags, which is proven to be an effective encoder of the properties inherited in complex data. Finally, the robustness of the proposed algorithm for FBM is demonstrated with the use of Monte Carlo simulations.

Keywords: anomalous diffusion; fractional Brownian motion; estimation; autocovariance function; neural network; Monte Carlo simulations

1. Introduction

Many single-particle tracking data related to the motion in crowded environments exhibit anomalous diffusion behavior [1,2]. This behavior is also visible in various phenomena such as finance [3,4], ecology [5], hydrology [6], and biology [7], as well as meteorology and geophysics [8,9]. Anomalous diffusion

behavior is manifested by deviations from the laws of Brownian motion (BM). One of the most common definitions of the anomalous diffusion process is expressed in the nonlinear behavior of its second moment:

$$\mathbb{E} \left[X^2(t) \right] \sim t^\alpha, \quad (1)$$

where α is the so-called anomalous diffusion exponent. When $\alpha = 1$, the process is classified as diffusion, while for $\alpha \neq 1$, we call it anomalous diffusion. More precisely, for $\alpha < 1$, it is called subdiffusion, while for $\alpha > 1$, it is superdiffusion. It should be mentioned that anomalous diffusion can be also related to the non-Gaussian probability density function of the corresponding process; for instance, see [10–13].

The class of anomalous diffusion processes is very rich. The most classical anomalous diffusion models are fractional Brownian motion (FBM) [8,14], Lévy stable motion [15], continuous-time random walk [16,17], and the subordinated processes (also called time-changed processes) [18–25]. We also mention here the processes with time- or position-dependent diffusion coefficients such as scaled Brownian motion [26,27] or heterogeneous diffusion models [28], as well as the superstatistical process [29] or diffusing diffusivity models (also called Brownian yet non-Gaussian diffusion process) [30]. We also refer the readers to the articles [31–34] and the references therein.

In the case of experimental data with anomalous dynamics, the main problem is first to recognize the type of anomaly and then to reconstruct properly the physical rules governing such a phenomenon. The main challenge is to identify the process from short trajectory inputs. Various approaches to address this problem can be found in the literature; for instance, see [35–42]. One of the simplest and most efficient approaches is based on the theoretical properties of the sample autocovariance function (ACVF) [43]. It is known that the ACVF is the characteristic that determines completely the centered Gaussian process. Thus, its sample version is a proper tool for the testing and estimation of the parameters of this process. The approach presented in the literature [43] is effective; however, it does not utilize all of the information contained in the sample ACVF for a given trajectory, i.e., only values of statistics for selected lags are utilized. Therefore, an evolution of this approach is proposed in this paper.

Herein, we compared three approaches that apply the ACVF for estimation of the anomalous diffusion exponent. The first one, the so-called *naive* method, uses the ACVF in one specific lag to estimate the anomalous diffusion exponent of the given process. The second algorithm is based on the ACVF corresponding to the vector of the selected lags. The last technique is based on the sample ACVF information extracted with the most informative subset of available lags. The designed novel method in the third approach is intuitive and it uses information directly available in a new fashion. Information retrieval from the sample ACVF vector is performed here with a learning-based scheme operating on the most informative lags, i.e., a feedforward neural network (FNN) [44] is designed and applied for solving the regression task. This approach has been proven to be an effective encoder of the properties inherited in complex data [45–47]. The goal is to preliminarily assess (using computer simulations) the predictive properties of an FNN for the estimation of the anomalous diffusion exponent based on a short data set. This exercise provides evidence of the performance of the simple version of the neural network (NN) adapted for the defined regimes (anomalous diffusion) and the ACVF vector, which is a projection of a valid complex and real trajectory, e.g., for particle movement in a solution. The reported results can be further enhanced by the exploitation of adaptive mechanisms inserted into recurrent neural networks (RNNs) and/or the detailed and proper inferences of multiscale pattern(s) for deep learning [1–5]. The advantage of the application of an FNN to the task defined above is that a trained neural network model can be a robust and efficient estimator for anomalous diffusion exponents, e.g., complex relations hidden in ACVF data can be extracted within one-step, concluding in a trained feedforward neural network.

The robustness of the introduced algorithm based on the ACVF and NN methods in comparison to the known ACVF-based techniques is demonstrated herein for the exemplary Gaussian process using Monte

Carlo simulations. We considered the FBM as an exemplary model with fractional dynamics; that is, the Gaussian process with stationary increments and the so-called self-similar property parametrized by the Hurst exponent $H = 0.5\alpha$, where α is the anomalous diffusion exponent given in (1).

The main goal of the paper was to prove that the incorporation of intelligent-based algorithms into classical estimation schemes can shed new light on the investigation of the anomalous diffusion phenomenon. Moreover, the classical tools enhanced by artificial intelligence (AI) methods are more effective in comparison to the known statistical algorithms used for anomalous diffusion parametrization.

The rest of the paper is organized as follows: In Section 2, we outline the definition of the fractional Brownian motion and the exemplary Gaussian process with anomalous diffusion behavior. Next, in Section 3, we discuss two of the estimation methods for the Hurst exponent based on the ACVF that are commonly used in various applications. In the next section, we present a new approach for the estimation of the H index. Namely, in the new algorithm, we combined the ACVF and NN methods. To demonstrate the effectiveness of the new approach, in Section 5, we present a simulation study where we compare the three considered algorithms for the estimation of the Hurst exponent. The last section concludes the paper and presents a future study.

2. Fractional Brownian Motion

Fractional Brownian motion (FBM) $\{X_H(t), t \geq 0\}$ with the Hurst index $H \in (0,1)$ is a continuous and centered Gaussian process defined through the following Langevin equation [14,48–50]:

$$\frac{dX_H(t)}{dt} = D\zeta_H(t), \tag{2}$$

where the parameter D is the diffusion coefficient. In Equation (2), $\{\zeta_H(t), t \geq 0\}$ is the fractional Gaussian noise process with the autocorrelation function satisfying the following:

$$\text{Corr}(\zeta_H(0), \zeta_H(t)) \sim 2H(2H - 1)Dt^{2(H-1)}, t \geq 0. \tag{3}$$

The FBM was introduced by Kolmogorov in 1940 (see [8]). FBM is the only Gaussian process with the self-similar property. Because the FBM is a centered Gaussian process, it can be also defined through the ACVF that, in this case, is given by [14]:

$$\mathbb{E}[X_H(t)X_H(s)] = \frac{1}{2}D \left(t^{2H} + s^{2H} - |t - s|^{2H} \right), \quad \text{where } t, s \geq 0. \tag{4}$$

Thus, for the given $t \geq 0$, $X_H(t) \sim \mathcal{N}(0, Dt^{2H})$. The FBM has stationary increments. Moreover, if $H > 1/2$, then the increments of the process are positively correlated, while for $H < 1/2$, they are negatively correlated. Moreover, for $H > 1/2$, the FBM exhibits the so-called long range dependence, which means that the following property is satisfied:

$$\sum_{n=1}^{\infty} \mathbb{E}[X_H(1)(X_H(n+1) - X_H(n))] = \infty. \tag{5}$$

As can be seen, for $H = 1/2$, the FBM reduces to the ordinary Brownian motion (BM). It should be mentioned that the FBM is considered one of the classical processes used to describe the anomalous diffusion phenomenon. Indeed, for $H < 1/2$, it exhibits subdiffusion behavior, while for $H > 1/2$, it shows superdiffusion behavior. To see the differences between the behavior of the trajectories corresponding to different anomalous types, in Figure 1, we demonstrate the exemplary trajectories of FBM for $H = 0.3$ (subdiffusion), $H = 0.5$ (diffusion), and $H = 0.7$ (superdiffusion).

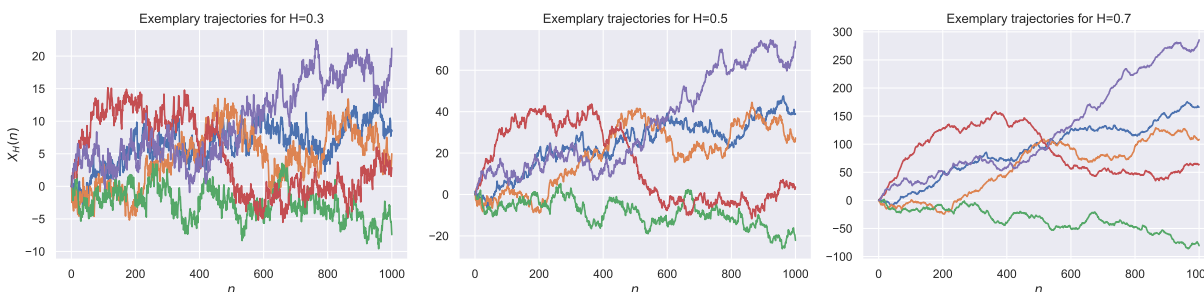


Figure 1. Exemplary trajectories for $H \in \{0.3, 0.5, 0.7\}$.

3. ACVF-Based Methods for the Estimation of the Hurst Exponent

In this article, we consider three approaches that utilize the ACVF in the estimation of the Hurst parameter H . Here, we depict two approaches known from the literature. The last technique is described in detail in the next section.

The methods presented in this section are based on the sample version of the ACVF. Let us consider the trajectory of FBM, $X_H = \{X_H(1), X_H(2), \dots, X_H(n)\}$, and the corresponding sample of increments, $\zeta_H = \{\zeta_H(1), \zeta_H(2), \dots, \zeta_H(n)\}$, where $\zeta_H(t) = X_H(t) - X_H(t - 1)$ for $t = 1, 2, \dots, n$. The sample ACVF for ζ_H is given by [51]:

$$\hat{\gamma}_\zeta(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} \zeta_H(t)\zeta_H(t + \tau), \quad \tau = 0, 1, \dots, n - 1. \tag{6}$$

The statistic $\hat{\gamma}_\zeta(\tau)$ is a rescaled estimator of the theoretical autocovariance function for $\gamma_\zeta(\tau)$, corresponding to lag τ , where $\gamma_\zeta(\tau) = \mathbb{E}(\zeta_H(1)\zeta_H(1 + \tau))$. One can easily show that the statistic (6) is a biased estimator of $\gamma_\zeta(\tau)$, namely:

$$\begin{aligned} \mathbb{E}[\hat{\gamma}_\zeta(\tau)] &= \frac{1}{n} \mathbb{E}\left[\sum_{t=1}^{n-\tau} \zeta_H(t)\zeta_H(t + \tau)\right] = \\ &= \frac{n - \tau}{n} \mathbb{E}[\zeta_H(1)\zeta_H(1 + \tau)] = \frac{D(n - \tau)}{2n} \left(|\tau + 1|^{2H} + |\tau - 1|^{2H} - 2|\tau|^{2H}\right) = \frac{n - \tau}{n} \gamma_\zeta(\tau). \end{aligned}$$

However, in our considerations, we used the biased version of the $\gamma_\zeta(\tau)$ estimator due to its lower variance in comparison to the unbiased one.

In the literature, a few possibilities of estimating the Hurst parameter H using the ACVF have been presented. In the simplest approach, we considered only the first lag $\tau = 1$, and compared the statistic $\hat{\gamma}_\zeta(1)$ with the desirable theoretical value of $\gamma_\zeta(1) = \frac{D}{2}(2^{2H} - 2)$. Thus, using the following relation:

$$\hat{\gamma}_\zeta(1) = \frac{D}{2} (2^{2H} - 2),$$

one can obtain the simple estimator of H :

$$\hat{H} = \frac{1}{2} \log_2 \left(\hat{\gamma}_\zeta(1) \frac{2}{D} + 2 \right). \tag{7}$$

As such an approach is simple, we refer to it as *naive* (the M1 method) in the following analyses. In this approach, the diffusion coefficient D is assumed to be known. However, in real applications, the D parameter can also be estimated as a sample variance of the vector ζ . As one can see, the ACVF for lag

$\tau = 1$ includes the necessary information about the FBM process. Unfortunately, if the corresponding measurements are burdened by an additive error (i.e., some measurement noise), $\gamma(1)$ changes accordingly, resulting in the need to consider more advanced techniques for the estimation of the H parameter; for instance, see [52,53] and the discussion therein.

In an alternative approach for estimating the H parameter, one can simultaneously use more lags τ . Thus, we can fit the function $\tau \mapsto \frac{D}{2} (|\tau + 1|^{2H} + |\tau - 1|^{2H} - 2|\tau|^{2H})$ to the empirical ACVF $\hat{\gamma}_\xi(\tau)$ for the corresponding lags $\tau = 1, 2, \dots, \tau_{max}$ in the least squares sense, i.e., the estimator is calculated as follows:

$$\hat{H} = \arg \min_{H \in (0,1)} \sum_{\tau=1}^{\tau_{max}} \left[\hat{\gamma}_\xi(\tau) - \frac{D}{2} (|\tau + 1|^{2H} + |\tau - 1|^{2H} - 2|\tau|^{2H}) \right]^2, \quad (8)$$

for some maximum lag τ_{max} . Again, if the diffusion coefficient is unknown, we can estimate it by considering $\arg \min$ over $(H, D) \in (0, 1) \times (0, \infty)$. This approach is unfortunately much more complex, as it requires nonlinear regression methods. In the further analysis, we refer to this approach as the M2 method.

4. ACVF and NN-Based Methods for the Estimation of the Hurst Exponent

As mentioned above, three methods for the estimation of the Hurst exponent H are compared in this article. All of the identification algorithms are based on the ACVF. The first (M1) and the second (M2) methods were described in the previous section. In this section, we present the algorithm based on the ACVF and NN methods, denoted as the M3 method in the simulation study.

One might expect that the two proposed methods provide the best estimation results when the data follows the “pure” theoretical model, i.e., FBM. In reality, this is rarely the case—often, the observed trajectories are biased by a measurement error and/or various interleaving processes cover the anomalous diffusion component in the acquired signal. A purely statistical approach, such as in the M1 and M2 methods, can bring about limitations in real-world data applications, whereas artificial intelligence has shown its potential to overcome parasitic conditions in data from numerous fields of applications [45,54–56]. This triggers applications of learning-based schemes that enable weakening of the initial assumptions related to data properties and model building, providing good estimators for a wide range of anomalous diffusion regimes (i.e., for a wide range of H values). The assumed architecture for the neural network model and the training process are crucial aspects for efficient data exploration in artificial intelligent schemes, with the latter being of particular importance for NN model performance [44]. As a consequence, the input data properties used for the NN model training condition the reliability of the observed outputs—i.e., the value of the estimated H and its uncertainty in the reported study.

In the previous section, we assumed that the relationship between the data and the estimated parameter is known and given by Equations (7) and (8) for the M1 and M2 methods, respectively. Now, we propose to use a feedforward neural network as the predictor of this theoretical relationship, i.e., $\mathbb{E}[H|\{\hat{\gamma}_\xi(\tau)\}]$. To be more precise, the FNN is proposed as the model of the hidden relationships in the experimental data. The last one means that obtaining a formal expression for the rules governing the phenomena in a real-world system is not the main subject of interest here, but encoded in the FNN model, these rules are used to enhance the reliability of the Hurst exponent estimation from the data that correspond to the FBM, according to our assumption. It is worth noting that the modeling of $\mathbb{E}[H|X_H]$ or $\mathbb{E}[H|\xi_H]$ is also possible with the NN-based approach; however, more sophisticated NN topologies are required to reconstruct the long dependency valid for the FBM model (for $H \gg 0.5$). Thus, dealing with long and varying input vector lengths is required to realize this task; models based on RNNs, long short-term memory (LSTM) neural networks [57], or other forms of intelligent recurrence should be used, which implies higher requirements regarding their training [58].

The information about the underlying process is concentrated in the first couple of lags of the sample ACVF. In this paper, 32 lags (including the 0th lag) were used as the input for the Hurst exponent estimation. In the next section, it is shown that this amount was sufficient for our study.

The proposed architecture of the neural network model consists of three hidden layers of consecutive sizes of 64, 64, and 32. The Swish-1 [59] activation function, defined as:

$$\text{Swish-}\beta(x) := x \times \sigma(\beta x) = \frac{x}{1 + e^{-\beta x}}, \quad (9)$$

was used for the neurons in each layer—in many applications, this expression outperforms the other activation functions [59].

In the designed FNN predictor, the size of the first layer was conditioned by the number of lags corresponding to the ACVF used during the experiment (i.e., 32 neurons were inserted into the first layer) and the output layer consisted of one neuron, which produced the H estimator as the model response. Since the estimated value of the Hurst exponent was within the range of $H \in (0, 1)$ and the FNN model designed for its estimation can produce all real numbers as the output, post-processing transformation needed to be applied to the response of the output neuron. Here, the sigmoid function, defined as:

$$\sigma(x) := \frac{1}{1 + e^{-x}}, \quad (10)$$

was used. The function given in (10) projects all real numbers to the interval $(0, 1)$.

To boost the FNN training process, the Adam optimization algorithm [60] was applied as it quickly converges to a minimum [61–64]. The mean squared error (MSE) was used to quantify the prediction error of the FNN model.

5. Simulation Study

The efficiency of the three methods (M1, M2, and M3) designed for H estimation is demonstrated in this section using computer simulations. Cholesky decomposition [65] was used for the generation of the FBM trajectories, since it allows to simulate outputs with extreme H parameter values (unlike Davies–Harte [66], which can fail to generate small samples for H parameters close to 1). Regarding the practical usefulness of the designed procedures, their efficiency can be expressed in the context of the length of the input trajectory required to estimate H with expected reliability. This study was performed for trajectories of various lengths, and the results are reported below.

The two statistical methods described in Section 3 were ready to use, whereas the designed feedforward neural network needed to be trained in advance, for which training and validation datasets were prepared.

The training dataset was formed with 1,572,64 FBM trajectories generated during computer simulations. This set consisted of vectors of different lengths (from 32 up to 1024) and referred to different H parameter values. For every trajectory length, $N \in \{32, 64, 128, 256, 512, 1024\}$, 262,144 trajectories were generated using computer simulations ($262,144 \times 6 = 1,572,864$), each with the H parameter selected randomly from the uniform distribution $\mathcal{U}(0, 1)$. Next, for each trajectory ζ_H , the ACVF (biased estimator (6), as introduced earlier) was calculated, resulting in a set of ACVFs of lengths N (N lags— $\tau \in \{0, 1, \dots, N - 1\}$) with corresponding H parameters. Using the same procedure, the validation and test subsets were generated, each of a size of 196,608 ($32,768 \times 6 = 196,608$).

The length of the ACVF vectors for the NN training was limited to 32 first lags (namely, $\tau \in \{0, 1, \dots, 31\}$). This selection was preceded by the calculation of MAE prediction error, such as in Figure 2—more input samples do not decrease the error, whereas a smaller input size increases the MAE value.

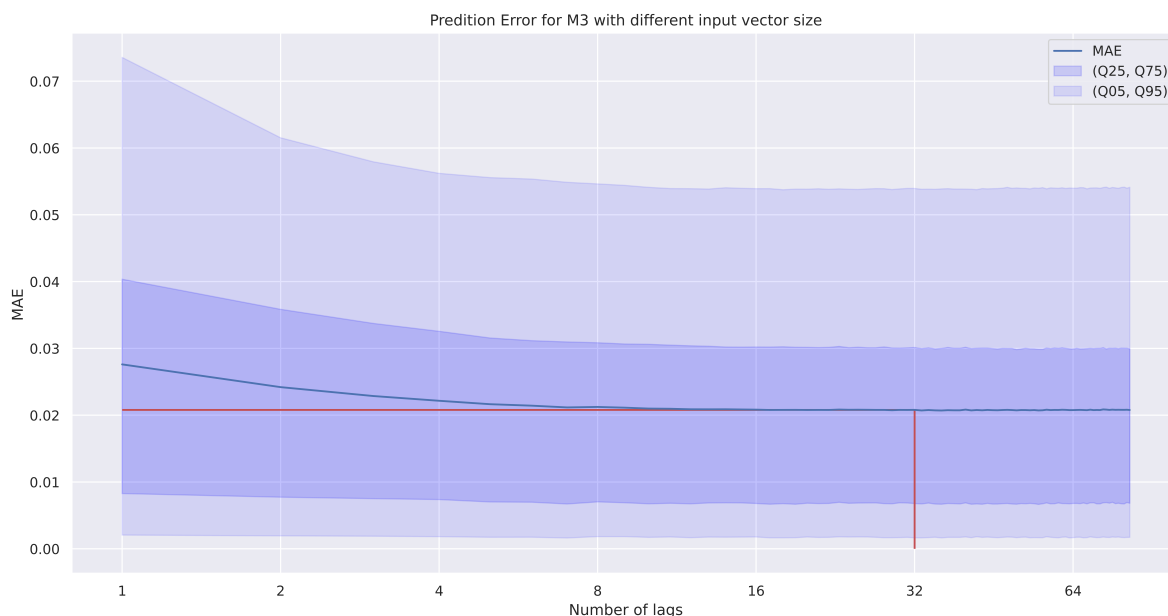


Figure 2. The MAE calculated for the M3 method when a different number of lags is used as the feedforward neural network (FNN) input (determining the size of the input layer of the FNN), depending also on the selected quantile; the number of lags (32) selected for use in the paper are marked with a red line.

To train the FNN (the architecture of the FNN as described in Section 4), input data were gathered into batches of 64 (the number of training examples used to calculate weight updates). The total number of NN parameters was 8385, trained for 13 epochs (the number of times that the training algorithm operated on the entire training dataset), for a total of $93 \text{ s} \times 13 \text{ epochs} \approx 20 \text{ min}$. The number of epochs was selected dynamically using the early stopping method [67]. Since the model did not improve the prediction error significantly after the third epoch, the training procedure could be stopped then (then, the training time would be 5 min). Calculations were performed on a PC with Intel Core i7 (3.7 GHz, 6 cores, 12 threads) and RAM of 64 GB.

The test dataset was used to compare the M1, M2, and M3 methods—for M1, the first lag was used ($\tau = 1$); for M2, the set of lags was $\tau \in \{0, 1, \dots, \tau_{max}\}$ (if not stated differently, $\tau_{max} = 31$); M3 always used 32 lags ($\tau \in \{0, 1, \dots, 31\}$). The metrics used were the absolute error and (for the aggregated results) the mean absolute error.

Figure 3 provides a comparison of the MAEs for the three methods (i.e., M1, M2, and M3) when dealing with trajectories of different lengths and with the diffusion characteristic (H parameter), grouped by the true parameter H into the bins $[0-0.2)$, $[0.2-0.4)$, $[0.4-0.6)$, $[0.6-0.8)$, and $[0.8-1)$.

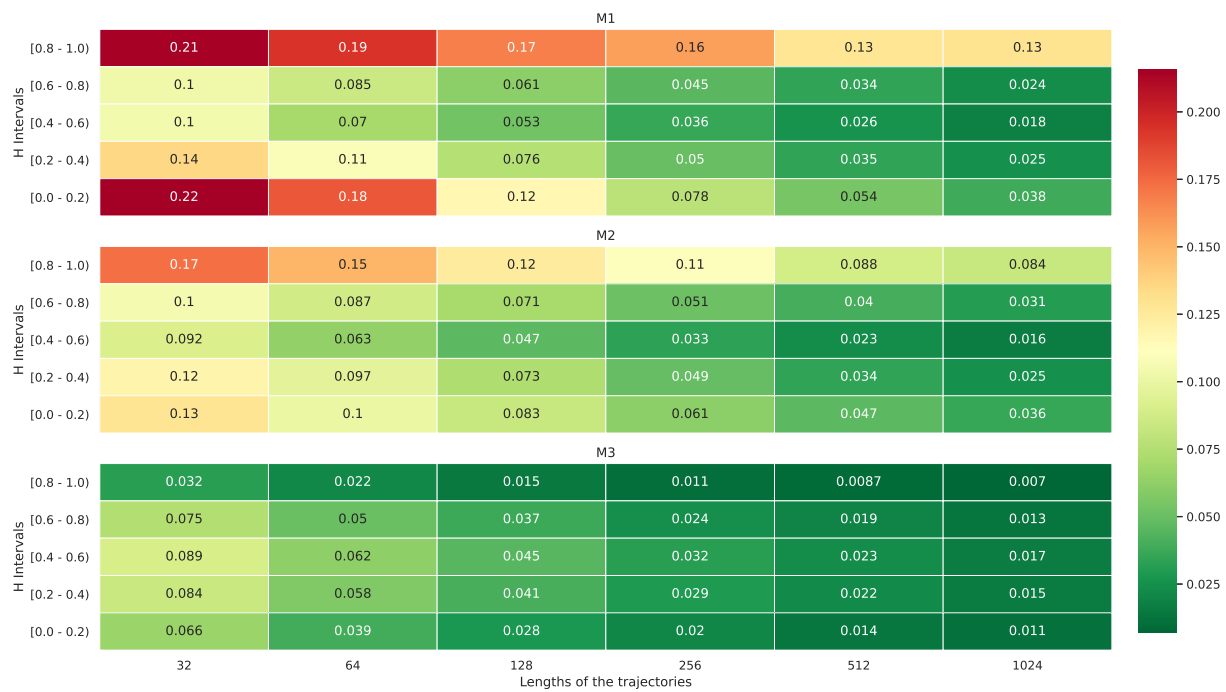


Figure 3. MAE heatmap for the M1, M2, and M3 methods depending on the length of the input trajectory and the value of the Hurst exponent applied during the computer simulations.

The conclusion is that the NN-based approach (M3) is more efficient than the other schemes studied in the paper, as it did not need long trajectories to estimate the Hurst exponent H with a minimized/minor error. For H close to 0 or 1, there was a significant difference in the MAEs between the considered methods—although the NN approach achieved a similar level of error to the other methods when fed samples of a size of 64 for the estimation, M1 and M2 struggled to deliver equivalent performance for longer inputs, i.e., up to 16-times longer trajectories, as was shown during the computer simulations. It is worth noting that the estimation error for the Hurst exponent in the diffusion case (i.e., $H \approx 0.5$) was similar for all of the considered approaches. In summary, since M1 estimated H reliably when using only the first lag, which was also true for the other methods, it was possible to distinguish between normal and anomalous diffusion using the information contained in the first lag or the first few lags.

To further compare these methods, the distribution of the (absolute) prediction error is shown in Figure 4. A logarithmic scale was applied to the y-axis in the boxplots to distinguish the performance of the following algorithms. The figure is divided into six parts, each reporting on the performance of the M1, M2, and M3 methods. In this way, it was possible to analyze how the distributions of the prediction error varied for each of these methods, the different lengths of the input trajectories, and also the different diffusion types (i.e., superdiffusion and subdiffusion); similarly to Figure 3, the input data were grouped evenly into five bins in reference to the true value of the H parameter.

Figure 4 depicts the spread of the estimation error and also reports on the number of outliers. The obtained results prove that the NN-based algorithm can more reliably estimate the H value for its following ranges (thus regimes of anomalous diffusion behavior) than in the two classical (M1 and M2) methods.

When it comes to the analysis of the multidimensional spread of the observed distributions, M3 performed similarly to the other methods. However, it is worth noting that M1 could not be used to obtain reliable results for some anomalous diffusion regimes, i.e., in several cases in Figure 4, the prediction

errors were over 0.5 (which means that M1 could not distinguish between sub- and superdiffusion in these cases). There were also some cases in the presented study when all of the methods working with the smallest considered sample size (32) struggled to distinguish between subdiffusion and superdiffusion. Nonetheless, M3 applied to two-times longer trajectories (i.e., 64 samples inserted as the input) performed reliably and efficiently through the whole range of H values.

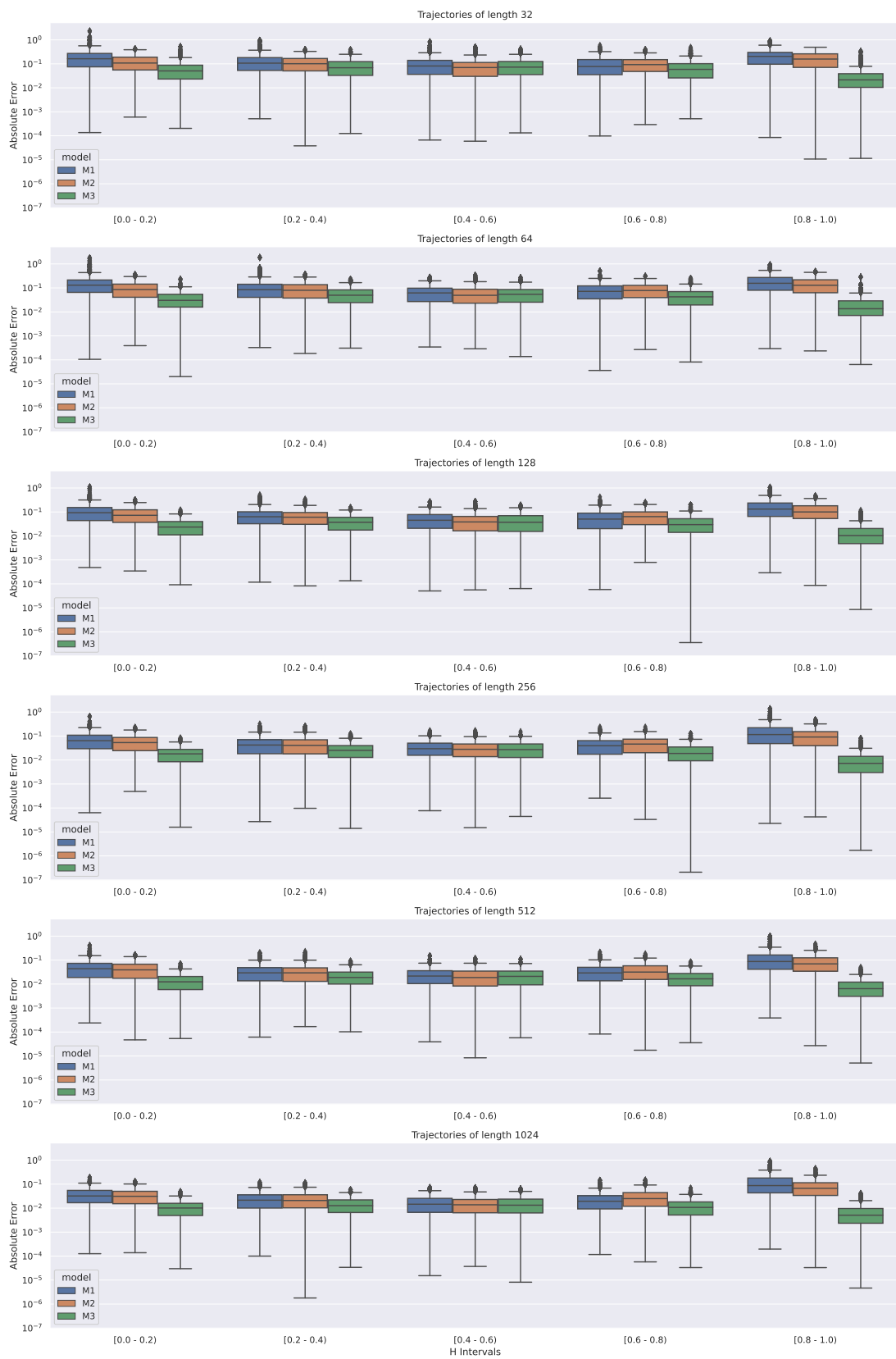


Figure 4. Absolute error calculated for the M1, M2, and M3 methods when fed with simulated input trajectories of different lengths and representing various modes of anomalous diffusion regimes (encoded with the H value).

It would be advantageous here to understand what makes M3 outperform the M2 algorithm. The most straightforward explanation relies on the fact that the input information contributes to outputs in various ways in M2 and M3. Namely, although M2 and M3 use the same input information, M2 explores each lag with the same importance, which is not the case for the FNN. The neural network-based method can learn inhomogeneous relationships contained in historical data, focusing on a specific subset of lags and leaving others with less influence on the observed output. This is exactly how the intelligent unpacking mechanism works in the M3 method.

This phenomenon can be clearly observed in Figure 5, where the MAE is compared for the methods working with trajectories of different lengths, for different H parameter values, and various numbers of lags used during the calculations. In the case of the M1 method, the results are presented only for $\tau = 1$, as there was no possibility to expand or shrink the set of lags here. M2 used lags up to τ_{max} for the Hurst exponent estimation, i.e., $(\tau \in \{0, 1, \dots, \tau_{max}\})$ with $\tau_{max} \in \{2, 3, \dots, 31\}$. The number of used lags could not be reduced for the M3 method since it required exactly 32 lags. To overcome this problem, a selected number of the first values of the calculated lags for the ACVF (this number is later referred to as non-zero lags) were extracted. The remaining lags were reset to zero during the calculations. This means that the relationships valid for the raw data were cut at the level of some lags, resembling the diffusion case $H = 0.5$ (i.e., there is no inter-dependency and the ACVF is equal to 0). All of these contributed to weakening the anomalous behavior (filling further lags with zero diminished/removed the long-range dependencies); thus, the reliability of the used methods for prolonged input data sequences could be improved, especially in the case of the M3 algorithm (see Figures 3 and 4).

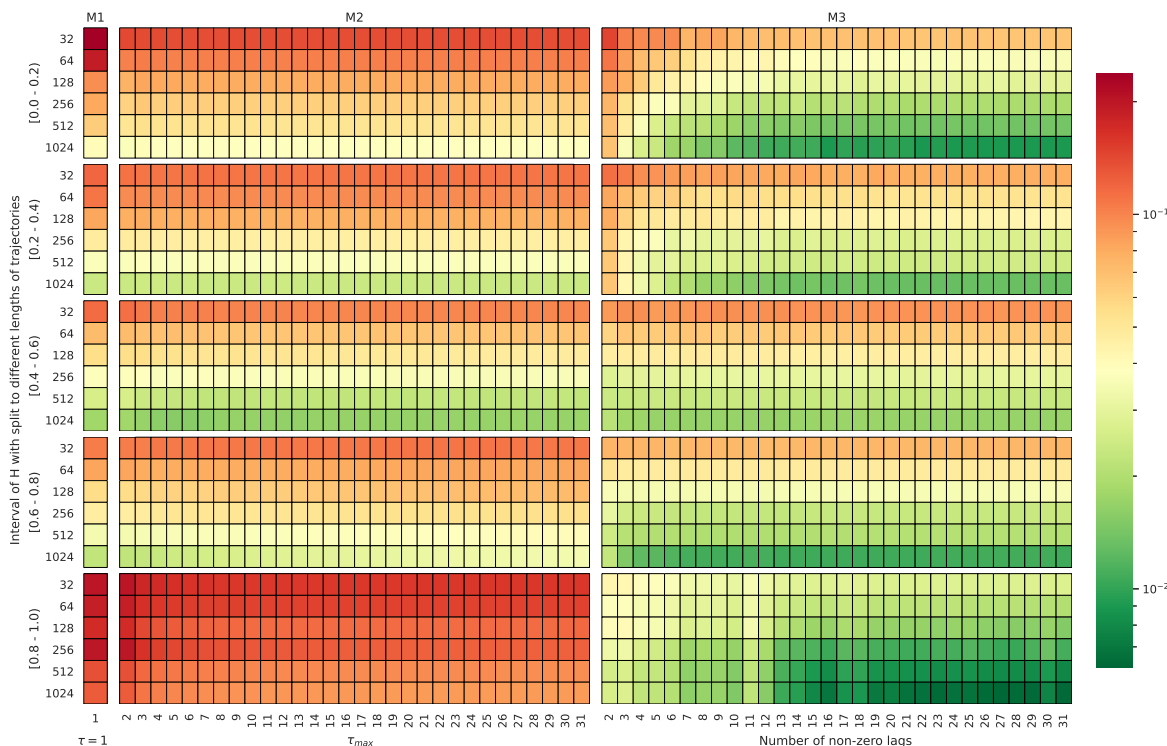


Figure 5. MAE heatmap calculated for the M1, M2, and M3 methods depending on the length of the input trajectory, τ_{max} , and for the following ranges of Hurst exponent values.

Figure 5 proves that using at least 20 lags was sufficient to minimize the error of H estimation in M3. However, the prediction error imperceptibly decreased for M2 with the addition of consecutive lags to the

input. One of the explanations for this observation might be that the magnitude of the contribution to the observed outputs (here, the value of the Hurst exponent) was smaller for the higher-order lags than for the first few lags. This means that the feedforward neural network more efficiently unpacked the information about the anomalous diffusion process than the M1 and M2 statistical schemes.

6. Summary and Conclusions

Anomalous diffusion is a complex phenomenon observed in physical systems. This complexity is inherited in recorded data, which, in practice, can be additionally corrupted by measurement noise. Moreover, anomalous diffusion components can emerge from a bunch of other regimes manifested in the observed system (and thus, in recorded data). Unpacking and disentangling the information contained in such data is a challenge, especially because complex interrelations are typically encoded in a small number of data samples. The statistical modeling of anomalous dynamics is quite well established and is concerned with the approximation of nontrivial patterns by the anomalous diffusion exponent. The problem is that for complex systems and the limited a priori knowledge available, the statistical inference can bring about non-robust estimation, especially when noisy components occur in data.

In recent years, there has been increasing interest in the use of AI methods of data exploration in various areas of science and engineering, especially when dealing with complex systems and a limited amount of preliminary information. In this paper, we proposed the application of a simple feedforward neural network to the quantification of anomalous dynamics. Namely, we compared this AI-based approach with statistical modeling, which allowed us to conclude that a combination of these two data exploration methods can decrease the estimation error of the anomalous diffusion exponent.

Herein, we considered the FBM, as it is one of the classical Gaussian processes that can be used for the description of anomalous diffusion behavior. Moreover, the FBM for special cases reduces to the classical BM, and thus is also useful for the analysis of diffusion processes.

Our approach is based on the ACVF, which completely describes the zero-mean Gaussian process. Thus, we selected the sample version of this statistic as a base for the estimation methodology. Via a simulation study, we proved that the classical approaches that utilize the sample ACVF are effective; however, when the process under consideration becomes anomalous diffusion, then their efficiency decreases. Thus, there is a need to incorporate more advanced techniques. In this article, we addressed such problems and proposed a simple modification of the classical ACVF-based approaches through the inclusion of NN-based methodology. Our simulations showed that the introduced estimation method outperformed the other considered approaches, especially for short trajectory lengths. This message is crucial for practical applications, where the real trajectories may be relatively short. In this case, the method based on a combination of ACVF and NN techniques seems to be more effective in contrast to the classical algorithms. It should be highlighted that although the presented approach in this paper is based on the ACVF, this methodology can be applied to any other statistic that is crucial for the estimation of anomalous diffusion processes, e.g., mean squared displacement, p -variation, and ergodicity breaking parameter. Moreover, the potential applicability of the described approach is much wider than the anomalous diffusion field. It could bring new insight into studies of physical systems with various properties reflected and decoded in the corresponding statistical quantities, i.e., long-range dependence, ergodicity breaking, or self-similarity. Additionally, the FBM was considered here only as an exemplary anomalous diffusion process, and the presented approach can be used for any other models with anomalous diffusion behavior. The designed method can also be used not only for estimation purposes, but also for the classification of the anomalous diffusion model governing physical phenomena. It could be directly achieved by the NN learning various ACVF formulas that uniquely define different Gaussian processes. Finally, outside of the Gaussian world, other appropriate statistics that evidently and precisely detect proper non-Gaussian

models can be effectively brought into the identification process. Thus, the introduced approach is universal in many areas and can be extended in various directions.

The combination of known statistical algorithms with the deep learning methodology is not a new idea in the area of anomalous diffusion phenomena analysis. Many authors have recognized that simple statistical methods in some cases seem to be inefficient for the proper identification or parametrization of the anomalous diffusion model, especially for recordings with short-length trajectories. In recent years, in the literature regarding anomalous diffusion processes, one can find the application of intelligent methods that enhance the classical approaches. In the physical sciences, deep learning methods have found very interesting applications. We mention here the new approaches based on the artificial NN algorithms (i.e., [68–73]) or the general machine learning methods applied to fractional dynamics analysis (i.e., [74–77]). However, to the best of our knowledge, a combination of ACVF- and NN-based methods has not been presented in the context of anomalous diffusion analysis.

In a future study, we plan to analyze the influence of additive noise on the estimation results for a model disturbed by external force. The same problem has been considered previously (e.g., in [53]), where the effectiveness of the time-averaged mean square displacement-based approach for the H parameter estimation was analyzed for the FBM with additive noise. Another future study will be related to the combination of the other time-averaged statistics [78] and deep learning methodology for the estimation problem in the anomalous diffusion regime.

Author Contributions: Conceptualization, G.S., I.J., A.W.; methodology, D.S., M.B., I.J., A.W.; software, D.S., G.S., M.B.; writing-original draft preparation, D.S., G.S., M.B., I.J., A.W.; writing-review and editing, D.S., G.S., M.B., I.J., A.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Nezhadhighi, M.G.; Rajabpour, M.A.; Rouhani, S. First-passage-time processes and subordinated Schramm-Loewner evolution. *Phys. Rev. E* **2011**, *84*, 011134. [[CrossRef](#)] [[PubMed](#)]
2. Failla, R.; Grigolini, P.; Ignaccolo, M.; Schwettmann, A. Random growth of interfaces as a subordinated process. *Phys. Rev. E* **2004**, *70*, 010101(R). [[CrossRef](#)] [[PubMed](#)]
3. Gabaix, X.; Gopikrishnan, P.; Plerou, V.; Stanley, H. A Theory of Power Law Distributions in Financial Market Fluctuations. *Nature* **2003**, *423*, 267–270. [[CrossRef](#)] [[PubMed](#)]
4. Ivanov, P.C.; Yuen, A.; Podobnik, B.; Lee, Y. Common scaling patterns in intertrade times of U.S. stocks. *Phys. Rev. E* **2004**, *69*, 05610. [[CrossRef](#)] [[PubMed](#)]
5. Scher, H.; Margolin, G.; Metzler, R.; Klafter, J. The dynamical foundation of fractal stream chemistry: The origin of extremely long retention times. *Geophys. Res. Lett.* **2002**, *29*, 1061. [[CrossRef](#)]
6. Doukhan, P.; Oppenheim, G. (Eds.) *Theory and Applications of Long-Range Dependence*; Birkhäuser Boston, Inc.: Boston, MA, USA, 2003.
7. Golding, I.; Cox, E.C. Physical Nature of Bacterial Cytoplasm. *Phys. Rev. Lett.* **2006**, *96*, 098102. [[CrossRef](#)]
8. Mandelbrot, B.B.; Ness, J.W.V. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Rev.* **1968**, *10*, 422–437. [[CrossRef](#)]
9. Stanislavsky, A.; Burnecki, K.; Magdziarz, M.; Weron, A.; Weron, K. FARIMA modelling of solar flare activity from empirical time series of soft X-Ray Solar emission. *Astrophys. J.* **2009**, *693*, 1877–1882. [[CrossRef](#)]
10. Zeng, C.; Chen, Y.; Yang, Q. The fBm-driven Ornstein-Uhlenbeck process: Probability density function and anomalous diffusion. *Fract. Calc. Appl. Anal.* **2012**, *15*, 479–492. [[CrossRef](#)]
11. Chen, W.; Sun, H.; Zhang, X.; Korošak, D. Anomalous diffusion modeling by fractal and fractional derivatives. *Comput. Math. Appl.* **2010**, *59*, 1754–1758. [[CrossRef](#)]

12. Mura, A.; Pagnini, G. Characterizations and simulations of a class of stochastic processes to model anomalous diffusion. *J. Phys. A* **2008**, *41*, 285003. [[CrossRef](#)]
13. Piryatinska, A.; Saichev, A.I.; Woyczynski, W.A. Models of anomalous diffusion: The subdiffusive case. *Phys. A Stat. Mech. Its Appl.* **2005**, *349*, 375–420. [[CrossRef](#)]
14. Beran, J. *Statistics for Long-Memory Processes*; Chapman & Hall: New York, NY, USA, 1994.
15. Samorodnitsky, G.; Taqqu, M. *Stable Non-Gaussian Random Processes*; Chapman & Hall: New York, NY, USA, 1994.
16. Metzler, R.; Klafter, J. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* **2000**, *339*, 1–77. [[CrossRef](#)]
17. Metzler, R.; Klafter, J. The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A* **2004**, *37*, R161–R208. [[CrossRef](#)]
18. Sato, K.I. *Lévy Processes and Infinitely Divisible Distributions*; Cambridge University Press: Cambridge, UK, 1999.
19. Wyłomańska, A. Arithmetic Brownian motion subordinated by tempered stable and inverse tempered stable processes. *Physica A* **2012**, *391*, 5685–5696. [[CrossRef](#)]
20. Wyłomańska, A. Tempered stable process with infinitely divisible inverse subordinators. *J. Stat. Mech.* **2013**, *10*, P10011. [[CrossRef](#)]
21. Wyłomańska, A.; Gajda, J. Stable continuous-time autoregressive process driven by stable subordinator. *Phys. A Stat. Mech. Its Appl.* **2016**, *444*, 1012–1026. [[CrossRef](#)]
22. Magdziarz, M.; Weron, A. Fractional Fokker-Planck dynamics: Stochastic representation and computer simulation. *Phys. Rev. E* **2007**, *75*, 056702. [[CrossRef](#)]
23. Magdziarz, M. Langevin Picture of Subdiffusion with Infinitely Divisible Waiting Times. *J. Stat. Phys.* **2009**, *135*, 763–772. [[CrossRef](#)]
24. Gajda, J.; Magdziarz, M. Fractional Fokker-Planck equation with tempered alpha-stable waiting times: Langevin picture and computer simulation. *Phys. Rev. E* **2010**, *82*, 011117. [[CrossRef](#)]
25. Gajda, J.; Wyłomańska, A. Fokker-Planck type equations associated with fractional Brownian motion controlled by infinitely divisible processes. *Physica A* **2014**, *405*, 104–113. [[CrossRef](#)]
26. Bodrova, A.S.; Chechkin, A.V.; Sokolov, I.M. Scaled Brownian motion with renewal resetting. *Phys. Rev. E* **2019**, *100*, 012120. [[CrossRef](#)] [[PubMed](#)]
27. Thalpa, S.; Wyłomańska, A.; Sikora, G.; Wagner, C.E.; Krapf, D.; Kantz, H.; Chechkin, A.V.; Metzler, R. Leveraging large-deviation statistics to decipher the stochastic properties of measured trajectories. *arXiv* **2020**, arXiv:2005.02099.
28. Fulinski, A. Anomalous Diffusion and Weak Nonergodicity. *Phys. Rev. E* **2011**, *83*, 061140. [[CrossRef](#)]
29. Beck, C.; Cohen, E.G.D. Superstatistics. *Physica A* **2003**, *322*, 267–275. [[CrossRef](#)]
30. Chechkin, A.V.; Seno, F.; Metzler, R.; Sokolov, I.M. Brownian yet Non-Gaussian Diffusion: From Superstatistics to Subordination of Diffusing Diffusivities. *Phys. Rev. X* **2017**, *7*, 021002. [[CrossRef](#)]
31. Sokolov, I.M. Models of anomalous diffusion in crowded environments. *Soft Matter* **2012**, *8*, 9043–9052. [[CrossRef](#)]
32. Klages, R.; Radons, G. (Eds.) *Anomalous Transport: Foundations and Applications*; Wiley: Hoboken, NJ, USA, 2008.
33. Hoefling, F.; Franosch, T. Anomalous transport in the crowded world of biological cells. *Rep. Prog. Phys.* **2013**, *76*, 046602. [[CrossRef](#)]
34. Metzler, R.; Jeon, J.; Cherstvy, A.; Barkai, E. Anomalous transport in the crowded world of biological cells. *Phys. Chem. Chem. Phys.* **2014**, *16*, 24128–24164. [[CrossRef](#)]
35. Sikora, G.; Burnecki, K.; Wyłomańska, A. Mean-squared displacement statistical test for fractional Brownian motion. *Phys. Rev. E* **2017**, *95*, 032110. [[CrossRef](#)]
36. Sikora, G.; Teuerle, M.; Wyłomańska, A.; Grebenkov, D. Statistical properties of the anomalous scaling exponent estimator based on time averaged mean square displacement. *Phys. Rev. E* **2017**, *96*, 022132. [[CrossRef](#)] [[PubMed](#)]
37. Taqqu, M.S.; Teverovsky, V.; Willinger, W. Estimators for Long-Range Dependence: An Empirical Study. *Fractals* **1995**, *3*, 785–798. [[CrossRef](#)]
38. Movahed, M.; Hermanis, E. Fractal analysis driver flow fluctuations. *Phys. A Stat. Mech. Its Appl.* **2008**, *387*, 915–932. [[CrossRef](#)]

39. Sikora, G. Statistical test for fractional Brownian motion based on detrending moving average algorithm. *Chaos Solitons Fractals* **2018**, *114*, 54–62. [[CrossRef](#)]
40. Carbone, A.; Kiyono, K. Detrending moving average algorithm: Frequency response and scaling performances. *Phys. Rev. E* **2016**, *93*, 063309. [[CrossRef](#)]
41. Arianos, S.; Carbone, A. Detrending moving average algorithm: A closed-form approximation of the scaling law. *Phys. A Stat. Mech. Its Appl.* **2017**, *382*, 9–15. [[CrossRef](#)]
42. Sikora, G.; Hoell, M.; Gajda, J.; Kantz, H.; Chechkin, A.; Wyłomańska, A. Probabilistic properties of detrended fluctuation analysis for Gaussian processes. *Phys. Rev. E* **2020**, *101*, 032114. [[CrossRef](#)]
43. Balcerek, M.; Burnecki, K. Testing of fractional Brownian motion in a noisy environment. *Chaos Solitons Fractals* **2020**, *140*, 110097. [[CrossRef](#)]
44. Bishop, C. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1996.
45. Jabłoński, I. Computer assessment of indirect insight during an airflow interrupter maneuver of breathing. *Comput. Meth. Progr. Biomed.* **2013**, *110*, 320–332. [[CrossRef](#)]
46. Kumar, R.; Srivastava, S.; Gupta, J.R.P.; Mohindru, A. Comparative study of neural networks for dynamic nonlinear systems identification. *Soft Comput.* **2019**, *21*, 103–114. [[CrossRef](#)]
47. Tutunji, T.A. Parametric system identification using neural networks. *Appl. Soft Comput.* **2016**, *47*, 251–261. [[CrossRef](#)]
48. Szymański, J.; Weiss, M. Elucidating the Origin of Anomalous Diffusion in Crowded Fluids. *Phys. Rev. Lett.* **2009**, *103*, 038102. [[CrossRef](#)] [[PubMed](#)]
49. Weiss, M. Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids. *Phys. Rev. E* **2013**, *88*, 010101. [[CrossRef](#)] [[PubMed](#)]
50. Krapf, D.; Lukat, N.; Marinari, E.; Metzler, R.; Oshanin, G.; Selhuber-Unkel, C.; Squarcini, A.; Stadler, L.; Weiss, M.; Xu, X. Spectral Content of a Single Non-Brownian Trajectory. *Phys. Rev. X* **2019**, *9*, 011019. [[CrossRef](#)]
51. Brockwell, P.J.; Davis, R.A. *Introduction to Time Series and Forecasting*; Springer: New York, NY, USA, 1994.
52. Sikora, G.; Kepten, E.; Weron, A.; Balcerek, M.; Burnecki, K. An efficient algorithm for extracting the magnitude of the measurement error for fractional dynamics. *Phys. Chem. Chem. Phys.* **2017**, *19*, 26566–26581. [[CrossRef](#)]
53. Lanoiselee, Y.; Grebenkov, D.; Sikora, G.; Grzesiek, A.; Wyłomańska, A. Optimal parameters for anomalous diffusion exponent estimation from noisy data. *Phys. Rev. E* **2018**, *98*, 062139. [[CrossRef](#)]
54. Borodinov, N.; Neumayer, S.; Kalinin, S.V.; Ovchinnikova, O.S.; Vasudevan, R.K.; Jesse, S. Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy. *Comput. Mater.* **2019**, *5*, 1–8. [[CrossRef](#)]
55. Gupta, S.; Gupta, A. Dealing with noise problem in machine learning data-sets: A systematic review. *Procedia Comput. Sci.* **2019**, *161*, 466–474. [[CrossRef](#)]
56. Gong, C.; Liu, T.; Yang, J.; Tao, D. Large-margin label-calibrated support vector machines for positive and unlabeled learning. *IEEE Trans. Neural Netw.* **2019**, *30*, 3471–3483. [[CrossRef](#)]
57. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
58. Pascanu, R.; Mikolov, T.; Bengio, Y. Understanding the exploding gradient problem. *CoRR* **2012**, *4*, 417.
59. Prajit, R.; Barret, Z.; Quoc, V. Searching for Activation Functions. *arXiv* **2017**, arXiv:1710.05941.
60. Diederik, P.; Jimmy, B. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
61. Lample, G.; Ballesteros, M.; Subramanian, S.; Kawakami, K.; Dyer, C. Neural architectures for named entity recognition. *arXiv* **2016**, arXiv:1603.01360.
62. Sabour, S.; Frosst, N.; Hinton, G. Dynamic Routing Between Capsules. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: New York, NY, USA, 2017; pp. 3856–3866.
63. Dosovitskiy, A.; Fischer, P.; Ilg, E.; Hausser, P.; Hazirbas, C.; Golkov, V.; van der Smagt, P.; Cremers, D.; Brox, T. FlowNet: Learning Optical Flow With Convolutional Networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 7–13 December 2015.

64. Mao, X.; Li, Q.; Xie, H.; Lau, R.; Wang, Z.; Smolley, S.P. Least Squares Generative Adversarial Networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
65. Watkins, D. *Fundamentals of Matrix Computations*, 2nd ed.; Pure and Applied Mathematics; Wiley-Interscience: New York, NY, USA, 2002.
66. Craigmile, P. Simulating a class of stationary Gaussian processes using the Davies—Harte algorithm, with application to long memory processes. *J. Time Ser. Anal.* **2003**, *24*, 505–511. [[CrossRef](#)]
67. Caruana, R.; Lawrence, S.; Giles, C. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In Proceedings of the Conference: Advances in Neural Information Processing Systems 13, Denver, CO, USA, 8–13 December 2000; pp. 402–408.
68. Kowalek, P.; Loch-Olszewska, H.; Szwabiński, J. Classification of diffusion modes in single-particle tracking data: Feature-based versus deep-learning approach. *Phys. Rev. E* **2019**, *100*, 032410. [[CrossRef](#)]
69. Janczura, J.; Kowalek, P.; Loch-Olszewska, H.; Szwabiński, J.; Weron, A. Classification of particle trajectories in living cells: Machine learning versus statistical testing hypothesis for fractional anomalous diffusion. *Phys. Rev. E* **2020**, *102*, 032402. [[CrossRef](#)]
70. Bondarenko1, A.N.; Bugueva, T.V.; Dedok, V.A. Inverse Problems of Anomalous Diffusion Theory: An Artificial Neural Network Approach. *J. Appl. Ind. Math.* **2016**, *3*, 311–321. [[CrossRef](#)]
71. Bosman, H.; Iacca, G.; Tejada, A.; Wortche, H.J.; Liotta, A. Inverse Problems of Anomalous Diffusion Theory: An Artificial Neural Network Approach. *Inf. Fusion* **2017**, *33*, 41–56. [[CrossRef](#)]
72. Dosset, P.; Rassam, P.; Fernandez, L.; Espenel, C.; Rubinstein, E.; Margeat, E.; Milhiet, P.E. Automatic detection of diffusion modes within biological membranes using backpropagation neural network. *BMC Bioinform.* **2016**, *17*, 197. [[CrossRef](#)]
73. Bo, S.; Schmidt, F.; Eichhorn, R.; Volpe, G. Measurement of anomalous diffusion using recurrent neural networks. *Phys. Rev. E* **2019**, *100*, 010102. [[CrossRef](#)] [[PubMed](#)]
74. Munoz-Gil, G.; Garcia-March, M.A.; Manzo, C.; Martin-Guerrero, J.; Lewenstein, M. Single trajectory characterization via machine learning. *New J. Phys.* **2020**, *22*, 013010. [[CrossRef](#)]
75. Arts, M.; Smal, I.; Paul, M.W.; Wyman, C.; Meijering, E. Particle Mobility Analysis Using Deep Learning and the Moment Scaling Spectrum. *Sci. Rep.* **2019**, *9*, 17160. [[CrossRef](#)] [[PubMed](#)]
76. Wagner, T.; Kroll, A.; Haramagatti, C.R.; Lipinski, H.G.; Wiemann, M. Classification and Segmentation of Nanoparticle Diffusion Trajectories in Cellular Micro Environments. *PLoS ONE* **2016**, *12*, e0170165. [[CrossRef](#)]
77. Granik, N.; Weiss, L.E.; Nehme, E.; Levin, M.; Chein, M.; Perlson, E.; Roichman, Y.; Shechtman, Y. Single-Particle Diffusion Characterization by Deep Learning. *Biophys. J.* **2019**, *117*, 185–192. [[CrossRef](#)]
78. Maraj, K.; Szarek, D.; Sikora, G.; Balcerek, M.; Wyłomański, A.; Jabłoński, I. Measurement instrumentation and selected signal processing techniques for anomalous diffusion analysis. *Meas. Sens.* **2020**. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).