



## Research Article

# Drug Treatment Effect Model Based on MODWT and Hawkes Self-Exciting Point Process

Xiaokai Nie <sup>1,2,3</sup> and Xin Zhao <sup>4</sup>

<sup>1</sup>School of Automation, Southeast University, Nanjing 210096, China

<sup>2</sup>Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing 210096, China

<sup>3</sup>Shenzhen Research Institute, Southeast University, Shenzhen 518057, China

<sup>4</sup>School of Mathematics, Southeast University, Nanjing 211189, China

Correspondence should be addressed to Xin Zhao; [mmxinzhao@hotmail.com](mailto:mmxinzhao@hotmail.com)

Received 8 August 2022; Revised 23 August 2022; Accepted 27 September 2022; Published 14 October 2022

Academic Editor: Sathishkumar V E

Copyright © 2022 Xiaokai Nie and Xin Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In precision medicine, especially in the pharmacodynamic area, the lack of an adequate long-term drug effect monitoring model leads to a quite low robustness to the instant drug treatment. Modelling the effect of drug based on the monitoring variables is essential to measure the drug benefit and its side effect preciously. In order to model the complex drug behavior in the context of time series, a sin function is selected to describe the basic trend of heart rate variable that is medically monitored. A Hawkes self-exciting point process model is chosen to describe the effect caused by multiple and sequential drug usage at different time points. The model considers the time lag between the drug given time and the drug effect during the whole drug emission period. A cumulative Gamma distribution is employed to describe the time lag effect. Simulation results demonstrate the established model effectively when describing the baseline trend and the drug effect with low noise levels, where the maximal overlap discrete wavelet transformation is utilized for the information decomposition in the frequency zone. The real data of the variables heart rate and drug liquemin from a medical database is analyzed. Instead of the original time series, scale variable  $s_4$  is selected according to the Granger cointegration test. The results show that the model accurately characterizes the cumulative drug effect with the Pearson correlation test value as 0.22, which is more significant for the value under 0.1. In the future, the model can be extended to more complicated scenarios through taking into account multiple monitoring variables and different kinds of drugs.

## 1. Introduction

Pharmacokinetics (PK) and pharmacodynamics (PD) are aimed at building mathematical models to extract scientific basis of modern pharmacotherapy. Specifically, pharmacokinetics describes the drug concentration-time courses in body fluids resulting from the intake of a certain dose of drug, and pharmacodynamics describes the observed effect arising from a certain drug concentration [1]. Between PK and PD, PD is more important for describing the variation of body conditions after drug treatment. Following a drug treatment, the original organ function can be enhanced or suppressed. Thus, determining the accurate dosage of the

drug is rather essential in the drug prescription. During the treatment process with a given drug dosage, the body condition may be improved but may also suffer some untoward reaction or adverse reaction, including side effect, toxic reaction, allergy reaction, secondary reaction, residual effect, and teratogenesis. Therefore, drug treatment can be helpful to the patient but might be harmful as well, and monitoring the drug effect by observing the clinically monitored variables plays an important role in understanding the mechanism.

The major objective of PD is to explore how the drug influences the monitored variables including heart rate, cardiac output, and mean arterial pressure. Modelling the drug

effect can help improve medical precision by applying more suitable treatment for the patient under a target criteria of the monitored variable. Drug effect modelling is critical to realize the online forecasting of the dynamic status of patient's disease. On the other hand, this can reduce the expenditure of medical resources including human resources like physicians and caregivers, drug usage, and economic cost.

Due to the special property of drug effect, the influence of the drug on patient is complex. The instant usage of drug has a long-term and dynamic effect on the patient. For example, there may exist a time delay for some drug to show the medical effects, which may be half an hour in some special circumstances. The time length taken by the drug effect is different not only among different drugs but also for different intake methods. For epinephrine, the effect of intramuscular injection is approximately maintained between 10 and 30 minutes, while the effect of subcutaneous injection can take as long as around 1 hour. Drug effect may increase rapidly and then decrease slowly afterwards before finally vanishes. Before the drug effect vanishes, more drug dosage may be required to be applied to the patient. Besides, multiple drug usages can cause multiple cumulative effects. Determining the correct drug dosage becomes quite challenging as the monitored variables vary with an unpredictable trend. To give a precise dose under a target value of the monitored variable, the whole past drug usage that is still effective should be taken into consideration, as their influence on the patient's future health condition still affects the current required dosage.

In this work, the influence of drug usage on the monitored variables is specifically analyzed, particularly in terms of the cumulative effect from the past drug usage. The rest of the paper is organized as follows: Section 2 gives the research review of drug effect modelling including the state of the art and our contributions, Section 3 presents the method proposed in this research, Sections 4 and 5 show the analysis and the simulation using the real medical data, and Section 6 presents a conclusion of the results and the viewpoints on the further research. All computations are implemented using the software R [2] of the version 4.0.2, and "waveslim" [3] was used for the wavelet decompositions. The hardware platform is iMac Pro (2017) configured with the processor 3.2 GHz 8-Core Intel Xeon W, the memory 32 GB 2666 MHz DDR4, and the Graphics Radeon Pro Vega 56 8 GB.

## 2. Literature Review

The approaches to establishing the models for characterizing the effect of drug on electrocardiogram (ECG) signals like heart rate can be generally divided into three main groups. The pros and cons of the available approaches are compared in detail in the following. The first group includes the typical statistical methods like linear regression, logit regression, analysis of variance (ANOVA) and basic statistical description methods like box plot. In order to study how the choice of anesthetic agent can greatly influence CSF tracer influx, Hablitz et al. [4] used the linear regression analysis. Linear

regression with extensions like Lasso, Ridge, and Elastic-net such penalized linear regression methods are good at explanation of the real application background and are simple to understand. Capel et al. [5] used 1- or 2-way ANOVA to investigate the propensity of hydroxychloroquine to cause bradycardia. ANOVA is effective to compare the drug efficiency difference among two or more comparison experiments. In the research of Sun et al. [6], pairwise network meta-analyze was performed using DerSimonian-Laird random effects model to analysis the impact of GLP-1 receptor agonists on blood pressure, heart rate, and hypertension among patients with type 2 diabetes. Gilbert and Krum [7] analyzed the effects of antihyperglycemic drug therapy on heart failure in diabetes by using the meta analysis as well. The pros and cons of this type of methods include that these methods are rather efficient when the data has simple format including cross section data, and the research purpose is more focused on producing the medical experiment results instead of the method improvement. When the data structure is complicated or the objective is lied on the methodology, more interests are concentrated on the subsequent second or third group of methods.

The second type of drug effect models mainly includes machine learning and deep learning methods such as ensemble decision tree methods (AdaBoost or XGBoost), neural networks (convolutional neural network and long-short term neural network), support vector machine, and Bayesian classifiers [8, 9]. For example, Sherman et al. [10] used machine learning methods to identify drug-cancer cell interaction based on some large in vitro databases. Bresso et al. [11] applied the methods decision trees and inductive logic programming to characterizing each drug's side-effect profiles in terms of drug and target properties. Costabal et al. [12] characterized the effect of 30 drugs on the QT interval using Gaussian process regression, sensitivity analysis, and uncertainty quantification. In the research of Juhola et al. [13], several machine learning methods are compared in the analysis of drug effects on iPSC cardiomyocytes, including decision trees,  $K$ -nearest neighbor nearest searching, multinomial logistic regression, and least squares support vector machines. Results show that random forests classifier and least-squares support vector machines have better performance than the other methods. Ekins et al. [14] compared the performance of different machine learning methods for end-to-end drug discovery and development including Naïve Bayesian, support vector machines, and more recently concerned deep neural networks. Madhukar et al. [15] proposed a Bayesian approach to the drug target identification using diverse data types. Soft computing techniques [16, 17], which although have not been directly applied to the drug effect modelling, can still be the promising solutions to the problem. The pros and cons of this type of methods include the following aspects. It is pretty impressive that the machine learning methods have high prediction accuracy and many of them have reliable generalization. However, most of the machine learning methods are based on quite complex parametric systems and lack of interpretability of parameter meanings, such that may not be the ideal tools for achieving both high accuracy and good explanation.

The third type of drug effect models mainly includes mathematical models such as Cox hazard model based on ordinary differential equations and partial differential equations. Chatterjee and Ahmad [18] applied the fractional-order differential equation model to analyzing the COVID-19 infection of epithelial cells. Huang et al. [19] gave the hierarchical Bayesian inference for HIV dynamic differential equation models which incorporated multiple treatment factors. Thirumalai et al. [20] proposed the fractional differential equations based method to analyze the combined drug therapy for HIV infection. Leander et al. [21] used the stochastic differential equations to analyze the mixed effects involved in pharmacokinetic data of nicotinic acid. Fuentes-Gari et al. [22] have compared the performance of different differential equations in the leukemia treatment, from which it can be found that stochastic differential mixed effects models are useful tools for identifying incomplete or inaccurate model dynamics. In general, those mathematical models have sound performance in both prediction accuracy and explanation.

The motivation of this research is that although the existing results have achieved good performance in PD, especially on drug effect modelling, they still have some shortcomings unsolved due to the complex background of pharmacodynamic, that is, the instant usage of drug can have long and complicated effect on the patient that is shown with the monitored variables. To illustrate this complicated effect, a stochastic modelling approach called Hawkes point process model is introduced to describe stochastic point process. In this work, a Hawkes point process model [23], namely, the self-exciting point process [24] is proposed to characterize the complex drug effect. The essential property of self-exciting point process is the occurrence of any event increased the probability of further events occurring which is consistent with the behavior of drug on patient. This model origins from the area of earthquake behavior analysis [25] and is recently extended to the fields of finance [26], disease prediction [27], and social media [28], where the model has shown satisfactory performance, but it has not been used for drug effect modelling yet.

The acquired data sets of the input variables in this work are time series that have clearly nonstationary property. They are also stochastic time series involved with random noise. Instead of using the original data in the model, the maximum overlap discrete wavelet transform (MODWT) is applied for the feather extraction [29]. MODWT is a method of decomposing the original time series into scaling and wavelet coefficients on different resolution levels, which can be seen as its smooth information and detail information [30]. In this research, these coefficients are used as the deep information involved in the original information. The method Fourier transform [31] can also decompose similar information but it is not suitable for non-stationary time series [32].

The innovation and contribution points in this research mainly have three aspects. The first one and also the main one is that self-exciting point process is proposed to describe the complex behavior of drug on patient. The second one is that MODWT is applied as the feather extraction method to

find deep information involved in the original time series. The third one is that the drug effect model developed in this research can describe the effect of a sequence of drug usage on the monitored variables instead of only one usage. The results can be further developed for assisting precision medicine.

### 3. Method

The data contain two variables, the drug usage variable  $d_i$  and the monitored variable  $x_t$ .  $d_i$  denotes the  $i$ th drug dosage at time  $t_i$ , and  $x_t$  gives the value of the monitored variable at time  $t$ . The goal is to describe  $x_t$  by using  $d_i$ , that is

$$\hat{x}_t = f(d_i, g), \quad (1)$$

where  $g$  represents the other information included in the prediction of  $x_t$ . In this paper, we mainly analyze the monitored variable heart rate as it is a typical variable that has similar properties to the other monitored variables. Variable  $x_t$  such as the changes of the heart rate mainly contains three components: the original waveform change involved in the heart beat change itself, the trend change resulting from the disease, and the trend change arising from the drug usage. In order to keep track of the trend caused by the drug and disease, the wavelet transform method MODWT is applied before implementing the self-exciting point process.

Since the wavelet basis Haar wavelet is simple to understand and has good modelling performance, it is chosen as the basis for the MODWT. The Haar scaling function is defined as

$$\phi(t) = \begin{cases} 1 & t \in [0, 1), \\ 0 & \text{else.} \end{cases} \quad (2)$$

Using dilation and translation, the scaling function at resolution level  $j$  and location  $k$  is

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k). \quad (3)$$

For more details, see Graps [33]. Then, the scale variables  $s_{j,k}$  can be given by

$$s_{j,k} = \langle x_t, \phi_{j,k} \rangle = \int_R x_t \phi_{j,k}(t) dt = 2^{j/2} \int_{I_{j,k}} x_t dt. \quad (4)$$

The Haar mother wavelet function  $\psi(t)$  is defined as

$$\psi(t) = \begin{cases} 1 & t \in [0, 0.5), \\ -1 & t \in [0.5, 1), \\ 0 & \text{else.} \end{cases} \quad (5)$$

And the wavelet function at resolution level  $j$  and location  $k$  is  $\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$ . The wavelet coefficients  $d_{j,k}$  are defined as  $d_{j,k} = \langle x(t), \psi_{j,k} \rangle$ . The scale variables vector  $s_j = (s_{j,0}, s_{j,1}, \dots, s_{j,t})$  and wavelet coefficient vector

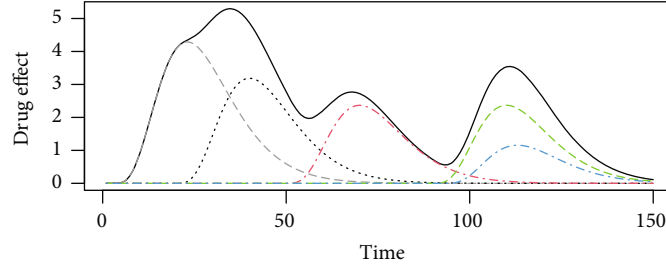


FIGURE 1: The drug effect of each drug entry. The black solid line represents the cumulative effect, and the dashed lines represent each drug entry.

TABLE 1: Parameter setting under different value.

Parameter	True parameter value	Parameter value range
$a_0$	70	[1, 100]
$b_0$	5	[0.1 10]
$\alpha_0$	1	[0.1 10]
$\beta_0$	2	[0.1 10]
$a_1$	1	[0.1, 50]
$b_1$	1	[0.01, 10]
$\kappa_1$	0.3	[0.01, 1]
$m_1$	-20	[-100, 98]
$a_2$	3	[0.01, 10]
$\kappa_2$	5	[2.1, 22]
$b_2$	0.4	[0.01, 2]

$d_j = (d_{j,0}, d_{j,1}, \dots, d_{j,t})$  become new variables which can be used for classification and regression [34, 35]. We name them as scale variables (or scale information) and detail variables (or detail information) in the following sections. For convenience,  $s_{j,t}$  is rewritten as  $s_{t,j}$ . By applying the Haar wavelet to the data  $x_t$ , the new data  $W_t$  is given by

$$W = \begin{bmatrix} d_{1,1} & \cdots & d_{1,J} & | & s_{1,1} & \cdots & s_{1,J} \\ d_{2,1} & \cdots & d_{2,J} & | & s_{2,1} & \cdots & s_{2,J} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ d_{T,1} & \cdots & d_{T,J} & | & s_{T,1} & \cdots & s_{T,J} \end{bmatrix}. \quad (6)$$

So far, the step of feature extraction has been finished by using MODWT with the input variables extracted as  $W$ . Instead of regarding all the variables in  $W$  as the target variables, Granger cointegration test is performed to verify how the drug influence the variables. The variable with the highest significance is chosen as the final target variable. In the Granger test, we set

$$x_t = \sum_{k=1}^m \alpha_k d_{t-k} + \sum_{k=1}^m \beta_k x_{t-k} + \varepsilon_t, \quad (7)$$

where  $\varepsilon_t$  is the random noise that follows the Gaussian distribution. Granger test is based on the  $F$  test, and the hypothesis is defined as

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0, \quad (8)$$

and the corresponding  $F$  statistic is

$$F = \frac{(\text{SSR}_r - \text{SSR}_{ur})/m}{\text{SSR}_{ur}/(n-k)}, \quad (9)$$

where  $\text{SSR}_r$  is the residual sum of squares being restricted, and  $\text{SSR}_{ur}$  is that unrestricted. If  $F > F_\alpha(m, n-k)$ , then, the hypothesis  $H_0$  that  $d_t$  is not the Granger reason of  $x_t$  is rejected.

By applying the Granger cointegration test to all the variables in  $W$ , the most significant variable can be selected as the final target variable, and we can regard the variable as the original  $x_t$  without the rest information  $g$ . In that case, the trend caused by the drug instead of  $x$  itself or the trend incurred by the disease can be discarded to some extent. For convenience, we still use the symbol  $x$  to represent the final target variable.

After getting the final target variable  $x$ , it can be applied to the Hawkes point process involved with the drug information. In the self-exciting point process, the target variable is a number that can be counted, and in this work, we extend it to a continuous context. As a result, the conditional intensity of  $x_t$  during the time interval  $(t_j, t_{j+1})$  is given by

$$\lambda_t = E_{t_j \leq t \leq t_{j+1}} [x_t - x_{t_j} | g_t], \quad (10)$$

where  $g_t$  includes the information available up to time  $t$ . After that, a fairly general self-exciting process can be defined in terms of an intensity of the form

$$\lambda_t = \mu_t + \sum_{t_i < t} \gamma(t - t_i, d_i), \quad (11)$$

where  $t_i$  is the time when the  $i$ th drug usage occurs.  $\mu_t > 0$  provides a base level for the process, and we set it as a sin function for simplicity,

$$\mu_t = a_0 + b_0 \sin(\alpha_0 t + \beta_0). \quad (12)$$

The function  $\gamma(t - t_i, d_i) \geq 0$  is defined as the exciting kernel of the process, which can be a Gamma distribution. The kernel provides the contribution to the intensity at time  $t$  that is made by all the previous drug usage events that

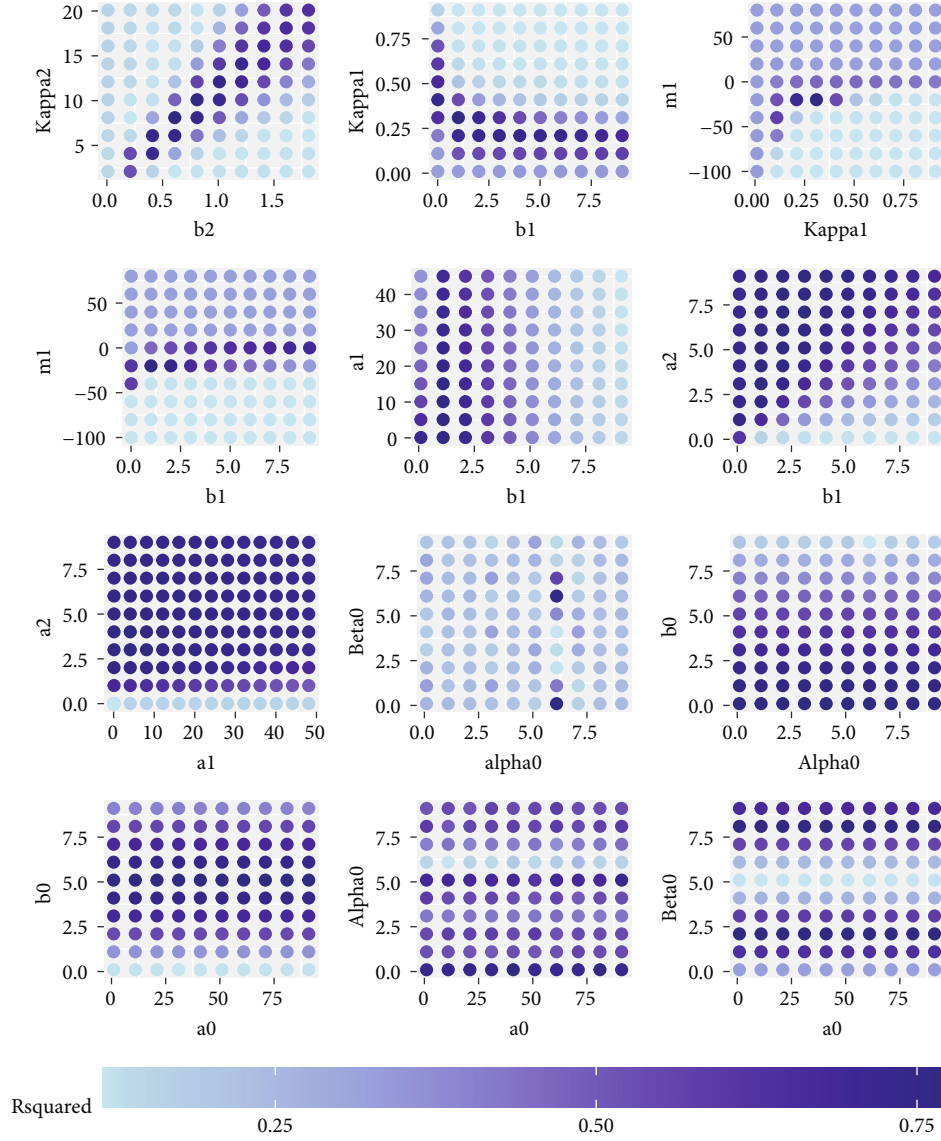


FIGURE 2: The  $R^2$  under different parameter settings. The true values are shown in Table 1.

occur at previous time  $t_i < t$ , with the associated drug dosage  $d_i$ . The meaning of the intensity function is that each event increases the intensity and then decays according to the function  $\gamma$  until the next drug usage occurs to push it up again. The chosen kernel function is a sigmoid function and a Gamma distribution function with the form as

$$\gamma(t - t_i, d_i) = \frac{1}{a_1 + b_1 e^{-\kappa_1(m_1 + d_i)}} \left\{ a_2 (t - t_i)^{\kappa_2 - 1} e^{-b_2(t - t_i)} \frac{b_2^{\kappa_2}}{\Gamma(\kappa_2)} \right\}, \quad (13)$$

where  $a_1$ ,  $a_2$ ,  $b_1$ ,  $b_2$ ,  $m_1$ ,  $\kappa_1$ , and  $\kappa_2$  are all parameters to be determined using the real data. In this kernel function, the sigmoid function is proposed to describe the instant drug effect, which we assume the appropriate drug dosage can bring about the corresponding correct effect while using an extra dosage can make little effect

when the drug has been already too much or rare. In that case, the sigmoid function is chosen to make it consistent with the assumptions. The Gamma function can be used to describe the trend of the drug effect from the increasing stage to the decreasing stage slowly. It can also describe the decaying process of the drug effect with the time. In the function of  $\lambda_t$ , the cumulative of the  $\gamma(t - t_i, d_i)$  describes all the previous drug effect until the current time. In order to calculate the parameters  $\theta = \{a_0, b_0, \alpha_0, \beta_0, a_1, a_2, b_1, b_2, m_1, \kappa_1, \kappa_2\}$ , the ordinary least squares method (OLS) is used to estimate these parameters. The target function to be minimized can be expressed as

$$L = \sum_{t=1}^T (x_t - \hat{x}_t)^2. \quad (14)$$

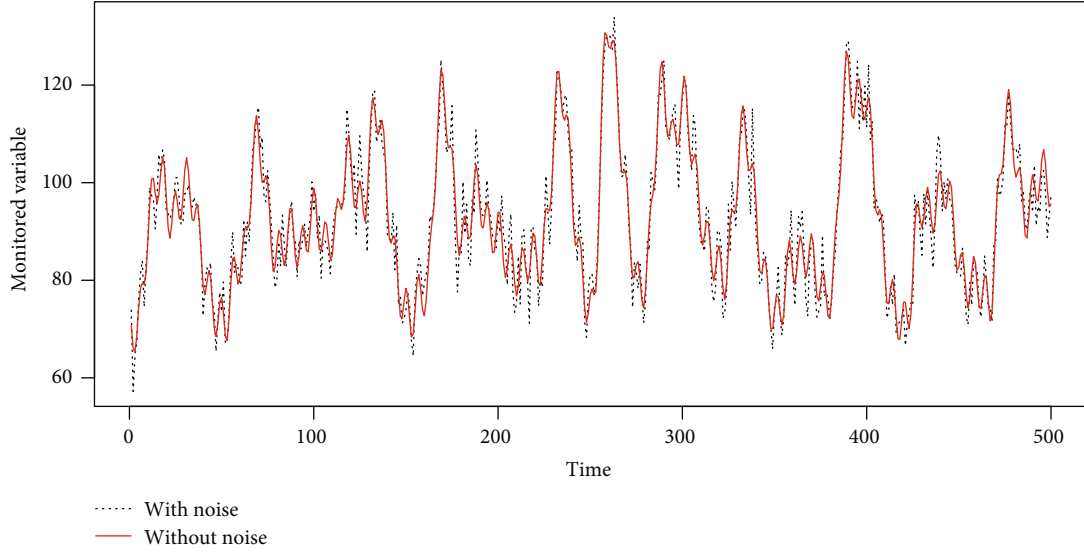


FIGURE 3: The simulated monitored variable with noise and the one without noise.

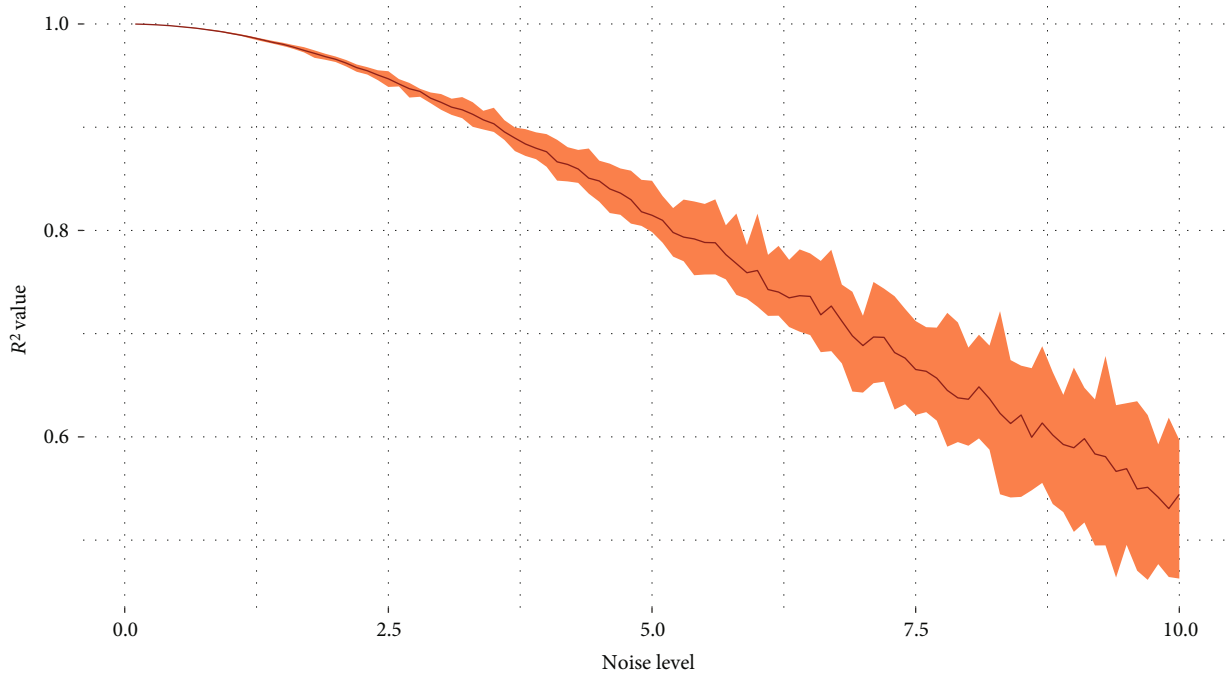


FIGURE 4: The best possible  $R^2$  under different noise levels.

Since  $\hat{x}_t$  is not easy to be obtained,  $\Delta\hat{x}_t$  is used for the replacement of it. Thus, the target function  $L$  can be given by

$$L \approx \sum_{t=1}^T (\Delta x_t - \Delta\hat{x}_t)^2. \quad (15)$$

Replacing  $\Delta\hat{x}_t$  with  $\lambda_t$  in  $L$  gives

$$L = \sum_{t=1}^T \left\{ \Delta x_t - \left( \mu_t + \sum_{t_i < t} \gamma(t - t_i, d_i) \right) \right\}^2. \quad (16)$$

It follows that

$$L = \sum_{t=1}^T \left\{ \Delta x_t - \left( \mu_t + \sum_{t_i < t} \frac{1}{a_1 + b_1 e^{-\kappa_1 d_i}} \left( a_2 (t - t_i)^{\kappa_2 - 1} e^{-b_2 (t - t_i)} \frac{b_2^{\kappa_2}}{\Gamma(\kappa_2)} \right) \right) \right\}^2. \quad (17)$$

Taking the derivative of each parameter in  $\theta$ , for parameter  $a_1$ , we have

$$a_1 : \sum_{t=1}^T \left\{ \Delta x_t - \hat{\mu}_t - \sum_{t_i < t} \gamma(t - t_i, d_i) \right\} \sum_{t_i < t} \frac{\partial \gamma}{\partial a_1}(t - t_i, d_i) = 0. \quad (18)$$

Similar results can be obtained for the other parameters in  $\theta$ . It is clear that we cannot get a closed form solution from the derivation as the estimated parameters are mutually dependent with each other. Hence, an optimization method of Broyden–Fletcher–Goldfarb–Shanno (BFGS) is proposed to minimize  $L$ , which is a quasi-Newton method, also known as the variable scale method. This algorithm improves the weakness of Newton’s method from the aspects that BFGS will not be easily affected by the initial value without the cost of computing the exact hessian matrix and its inverse in the process of each step of optimization. The NFGS, namely, the Newton improved BFGS, has the characteristic of fast searching of Newton’s method and has an improvement in efficient searching for the global optimal solution.

The measurement of the goodness of fit has the metric using  $R^2$ , which ranges from 0 to 1, with 1 as the best fit.  $R^2$  has the form as

$$R^2 = \frac{SST - SSE}{SST}, \quad (19)$$

where SST is the total sum of squares, and SSE is the residual (error) sum of squares,

$$SST = \sum_{t=1}^T (\Delta x_t - \Delta \bar{x}_t)^2, \quad (20)$$

$$SSE = \sum_{t=1}^T (\Delta x_t - \Delta \hat{x}_t)^2. \quad (21)$$

In addition to the  $R^2$ , the Pearson correlation test value is also selected for comparison, which is formulated as

$$\text{corr} = \sqrt{\frac{SST - SSE}{SST}} = R. \quad (22)$$

Some other evaluation criteria to assess the performance of the models include mean absolute error and mean percentage error. They all give the measurement of the model fitness, and  $R^2$  and the correlation test can have similar or better effect in measuring the results as their values are constrained in 0 to 1 or -1 to 1, such that the performance can be compared with the significant level. In conclusion, we develop a method to analysis the influence of drug usage on the monitored variable by using self-exciting point process with the kernel function of Gamma distribution.

#### 4. Simulation

In this section, we analyze the performance of the method under different noise levels. Since the drug effect can produce the cumulative effect, as shown in Figure 1, we will also analyze how the drug effect influences the model performance. In this Figure 1, the black line represents the cumulative effect  $\sum_{t-t_i} \gamma(t-t_i, d_i)$ , and the dashed lines represent the separate drug effect  $\gamma(t-t_i, d_i)$  for each  $i$ .

In order to explore how the parameters influence the model performance, the simulation is performed using dif-

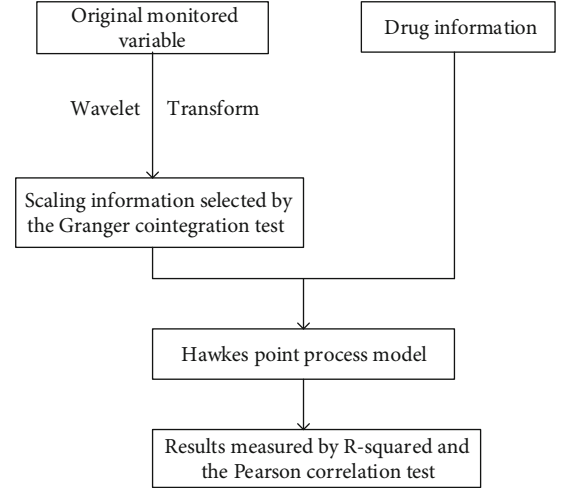


FIGURE 5: The flowchart of the proposed method.

ferent parameter settings. The true parameter and the simulated parameters are shown in Table 1. With the parameter given in Table 1, the  $R^2$  results are shown in Figure 2. It can be seen that some parameters have few influence on the  $R^2$ , including  $b_1, a_1, a_2, b_0, a_0$ . The parameters that have the apparent influence on the  $R^2$  include  $b_2, \kappa_2, \kappa_1, m_1, b_0, \alpha_0, \beta_0$ . From Figure 2, it can be seen that when the values of  $b_2$  and  $\kappa_2$  keep in a fixed ratio around 10, the  $R^2$  keeps in its good performance. When  $\kappa_1$  increases from 0 to around its best value 0.3, the  $R^2$  increases, but decreases afterwards when  $\kappa_1$  continues to grow. For the parameter  $m_1$ ,  $R^2$  has the similar trend. The model performance is not good when  $m_1$  is below its best value but performs relatively well when  $m_1$  is above its best value. For  $b_0, \alpha_0$  and  $\beta_0$ , as they are in the sin function,  $R^2$  shows some kind of routinely performance.

The level of noise also influences the best fitting ability. That is why MODWT is applied before the model training. For example, as shown in Figure 3, by adding the noise which follows the Gaussian distribution, the monitored variable with noise becomes far away from the one without noise. This gap causes the  $R^2$  to be smaller than it should be. By setting the noise level as 0.1 to 10, we can get the best possible  $R^2$  value by choosing the best parameters, and the results are shown in Figure 4. The  $R^2$  decreases as the noise level increases. When the variance of the noise level is bigger than 4, the  $R^2$  decreases below 0.9. When the noise level is 5, the  $R^2$  decreases below 0.8. This result also shows the best possible  $R^2$  under different noise levels.

#### 5. Real Data Analysis

The real data we use in this work is the circulatory failure data Hyland et al. [36]. The original source of the data comes from the Medical Information Mart for Intensive Care-(MIMIC-) III database [37], which provides the critical care data of over 40,000 patients admitted to intensive care units at the Beth Israel Deaconess Medical Center. Importantly, MIMIC-III was deidentified, and patient identifiers were

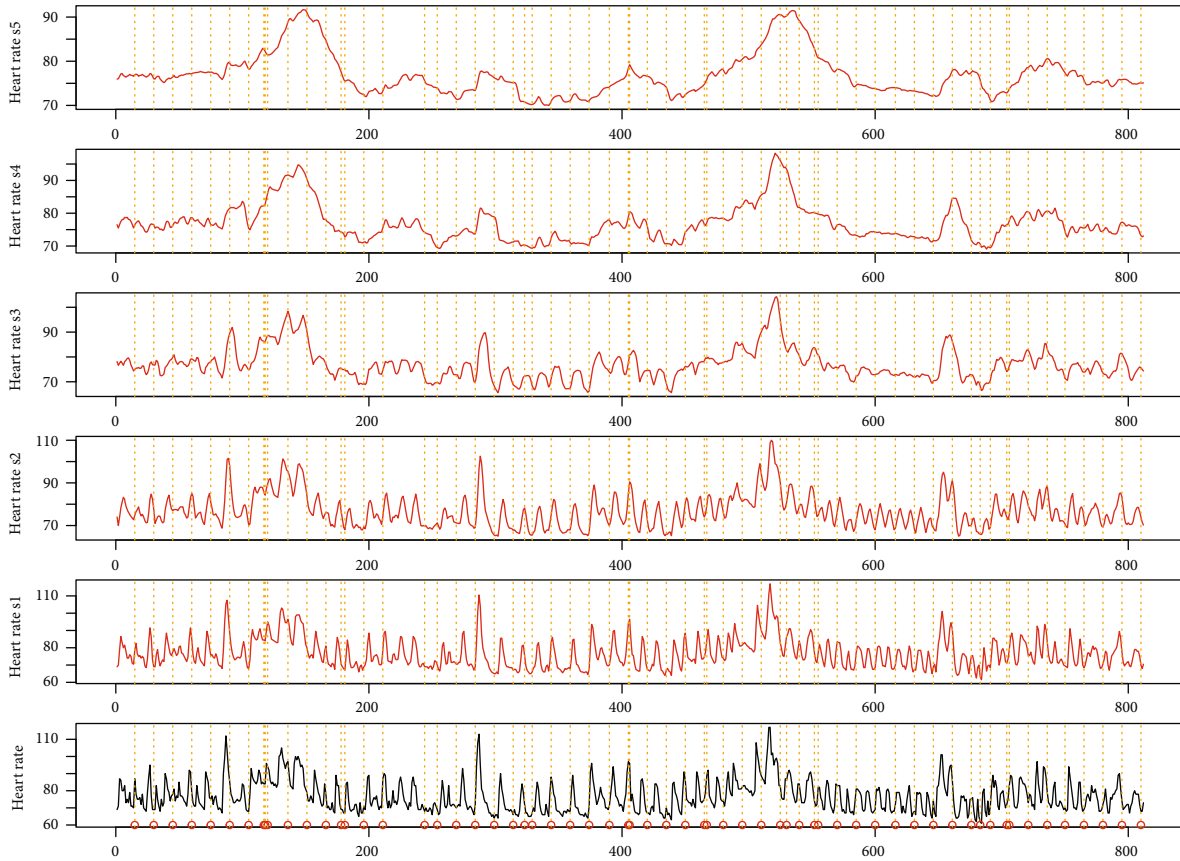


FIGURE 6: The scale variables of the original heart rate time series with drug labelled as the vehicle lines.

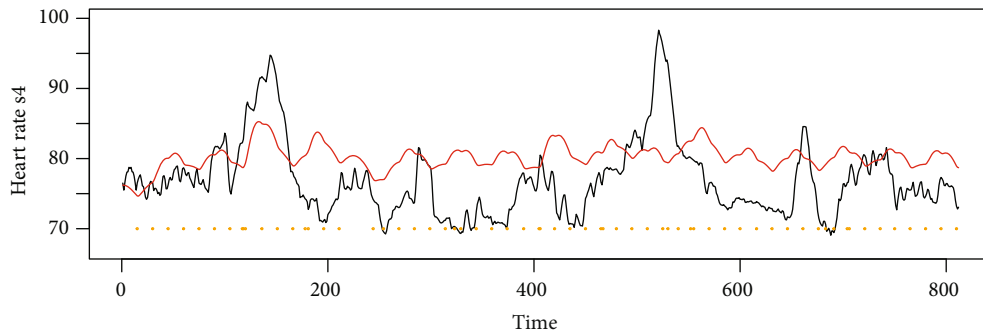


FIGURE 7: The scale variables of the original heart rate time series with drug labelled as the vehicle lines.

removed according to the Health Insurance Portability and Accountability Act Safe Harbor provision. MIMIC-III has been integral in driving large amounts of research in clinical informatics, epidemiology, and machine learning. The data we use contain 812 observations, with 1-minute gap per observation as the resolution level. There are totally 18 variables in the data including heart rate, MAP, cardiac output, and SpO<sub>2</sub>.

Instead of using all the variables, the variable heart rate is selected as the monitored variable as it has quite few missing values. The minimum and maximum of the heart rate are 61 and 117, respectively. The average and standard deviation are 77 and 9.17, respectively. The drug data in use is the amount of dosage of drug liquemin with unit 5000 U/ml

for per dosage. The dosage in use in this real data is 25. The number of usage is 63 times. Liquemin is used to decrease the clotting ability of the blood and help prevent harmful clots from forming in blood vessels. This medicine is sometimes called a blood thinner, although it does not actually thin the blood. Specifically, it is also used in the treatment of heart attacks and unstable angina. Two time-series phases are selected for the real data analysis which contain only liquemin. The flowchart of the proposed method is shown in Figure 5. As shown in Figure 6, the original heart rate variable is transformed by using MODWT to leave out the noise, and the scaling information is selected by using the Granger cointegration test. After that, the scaling information and the drug information are put into the



Hawkes point process model with the  $R^2$  or correlation test as the results. In this case, scale variable  $s_4$  on resolution level is selected as the target variable  $x$  according to the Granger cointegration test, and the results are shown in Figure 7. The computed parameters are  $\theta = (75, 1.21, 0.13, 1.39, 0.30, 2.08, 0.80, 0.47, 0.0022, 2.77, 0.12)$ . The  $R^2$  is not big as there are other trend inside the original data, but the correlation is obvious with a Pearson test value as 0.22.

Instead of comparing to the time series of the original data, the scale variables on level 4 is compared with the model output since it leaves out most of the noise and with only the trend left for comparison. From Figure 7, we can see that most of the increases and drops are caught by the red line, which means the proposed Hawkes model can effectively measure the drug effect. The Pearson test result is 0.22 that is significantly below the level of 0.1. The obtained model is given by

$$\begin{aligned} \lambda_t = & 75 + 1.21 \sin(0.13t + 1.39) \\ & + \sum_{t_i < t} \frac{1}{0.3 + 2.08e^{-0.8(-0.47+d_i)}} \\ & \cdot \left\{ 0.0022(t - t_i)^{1.77} e^{-0.12(t-t_i)} \frac{0.12^{2.77}}{\Gamma(2.77)} \right\}. \end{aligned} \quad (23)$$

To sum up, due to the complex development of the monitored variables, there are still some trends caused by the disease that cannot be fully described by our model. We propose a model that gives a different and efficient method for the drug effect measurement in pharmacodynamic.

## 6. Conclusion

In this research, we develop a Hawkes model by using self-exciting point process and sin function to describe the development of medical monitored variable heart rate. Self-exciting point process can describe the effect of the drug and the dosage of the drug. The sin function can describe the basic trend of heart rate. By combining these two functions together, some of the heart rate trend can be described. The model can be used for drug effect prediction and solve the problem of drug precision suggestion. It can also be used for the medical assistant by giving more precise drug dosage description. Specifically, the results show that the model can have a correlation Pearson test to be significant under the significant level of 0.1. The increases and decreases of the trained drug effect can fit the trend of the real drug effect.

The limitations of this research are summarized as the following. In our research, only one kind of drug is modelled, but if many kinds of drug are considered in the model, the model can also be complex. Specifically, in the further research, more complex model can be developed to describe the details of the monitored variable trend including the trend caused by the disease, in addition to the basic trend and drug effect trend. The trend caused by the disease is quite complex to be described but still can be modelled. The analysis of multiple drugs is not simply based on the adding of single drugs, in that case, using many self-

exciting point process cannot meet the demand of multiple drug effect analysis. Mutual-exciting point process Hawkes model can be considered to describe the effect of multiple kinds of drug and cumulative drug effect dosage. If multiple drug effect can be modelled, then the further analysis can be concentrated on the inverse problem of drug dosage decision under a predefined monitored variable value.

The proposed modelling solution can be further extended to some other areas in addition to the medical area. For example, in the financial area, the model can be used to monitor the effect of economic events or political events to predict the stock price and security price trend after the event occurs. In the environmental area, the model can be used to monitor the effect of air pollution by decomposing the original time series into three aspects, namely, the basic routine trend, the trend arising from events, and that caused by solar circles. In conclusion, the model developed in this research is an efficient and generalized model for time series data which contains multiple kinds of trends either embraced by the time series itself or caused by the external events.

## Abbreviations

MODWT:	Maximal overlap discrete wavelet transform
PK:	Pharmacokinetic
PD:	Pharmacodynamic
ECG:	Electrocardiogram
ANOVA:	Analysis of variance
BFGS:	Broyden–Fletcher–Goldfarb–Shanno
MIMIC:	Medical Information Mart for Intensive Care.

## Data Availability

The source code in the method are available from the corresponding author upon request. The real data in application can be requested from Hyland et al. [36].

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

Xin Zhao and Xiaokai Nie gratefully acknowledge the support from the Fundamental Research Funds for the Central Universities (2242020R40073, 2242022k30038, MCCSE2021B02, and 2242020R10053), Southeast University Zhishan Youth Scholar Foundation, Natural Science Foundation of Jiangsu Province (BK20200347 and BK20210218), Nanjing Scientific and Technological Innovation Foundation for Selected Returned Overseas Chinese Scholars (1108000241), Jiangsu Foundation for Innovative and Entrepreneurial Doctor (1107010306 and 1108000245), Guangdong Basic and Applied Basic Research Foundation (2020A1515110129), and National Natural Science Foundation of China (62103105, 12201108, and 12171085).

## References

- [1] B. Meibohm and H. Derendorf, "Basic concepts of pharmacokinetic/pharmacodynamic (pk/pd) modelling," *International Journal of Clinical Pharmacology and Therapeutics*, vol. 35, no. 10, pp. 401–413, 1997.
- [2] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [3] B. Whitcher, *Waveslim: Basic Wavelet Routines for One-, Two- and Three-Dimensional Signal Processing. R Package Version 1.7.5.1*, CRAN, 2019.
- [4] L. M. Hablitz, H. S. Vinitzky, Q. Sun et al., "Increased glymphatic influx is correlated with high eeg delta power and low heart rate in mice under anesthesia," *Science Advances*, vol. 5, no. 2, p. eaav5447, 2019.
- [5] R. A. Capel, N. Herring, M. Kalla et al., "Hydroxychloroquine reduces heart rate by modulating the hyperpolarization-activated current  $I_f$ : novel electrophysiological insights and therapeutic potential," *Heart Rhythm*, vol. 12, no. 10, pp. 2186–2194, 2015.
- [6] F. Sun, S. Wu, S. Guo et al., "Impact of glp-1 receptor agonists on blood pressure, heart rate and hypertension among patients with type 2 diabetes: a systematic review and network meta-analysis," *Diabetes Research and Clinical Practice*, vol. 110, no. 1, pp. 26–37, 2015.
- [7] R. E. Gilbert and H. Krum, "Heart failure in diabetes: effects of anti-hyperglycaemic drug therapy," *The Lancet*, vol. 385, no. 9982, pp. 2107–2117, 2015.
- [8] H. Ding, I. Takigawa, H. Mamitsuka, and S. Zhu, "Similarity-based machine learning methods for predicting drug–target interactions: a brief review," *Briefings in Bioinformatics*, vol. 15, no. 5, pp. 734–747, 2014.
- [9] N. Stephenson, E. Shane, J. Chase et al., "Survey of machine learning techniques in drug discovery," *Current Drug Metabolism*, vol. 20, no. 3, pp. 185–193, 2019.
- [10] J. Sherman, G. Verstandig, J. W. Rowe, and Y. Brumer, "Application of machine learning to large in vitro databases to identify drug-cancer cell interactions: azithromycin and KLK6 mutation status," *Oncogene*, vol. 40, no. 21, pp. 3766–3770, 2021.
- [11] E. Bresso, R. Grisoni, G. Marchetti et al., "Integrative relational machine-learning for understanding drug side-effect profiles," *BMC Bioinformatics*, vol. 14, no. 1, pp. 1–11, 2013.
- [12] F. S. Costabal, K. Matsuno, J. Yao, P. Perdikaris, and E. Kuhl, "Machine learning in drug development: characterizing the effect of 30 drugs on the qt interval using gaussian process regression, sensitivity analysis, and uncertainty quantification," *Computer Methods in Applied Mechanics and Engineering*, vol. 348, pp. 313–333, 2019.
- [13] M. Juhola, K. Penttinen, H. Joutsijoki, and K. Aalto-Setälä, "Analysis of drug effects on ipsc cardiomyocytes with machine learning," *Annals of Biomedical Engineering*, vol. 49, no. 1, pp. 129–138, 2021.
- [14] S. Ekins, A. C. Puhl, K. M. Zorn et al., "Exploiting machine learning for end-to-end drug discovery and development," *Nature Materials*, vol. 18, no. 5, pp. 435–441, 2019.
- [15] N. S. Madhukar, P. K. Khade, L. Huang et al., "A bayesian machine learning approach for drug target identification using diverse data types," *Nature Communications*, vol. 10, no. 1, pp. 1–14, 2019.
- [16] T. Hempel and A. Al-Hamadi, "An online semantic mapping system for extending and enhancing visual slam," *Engineering Applications of Artificial Intelligence*, vol. 111, p. 104830, 2022.
- [17] J. Ma, D. Xia, H. Guo et al., "Metaheuristic based support vector regression for landslide displacement prediction: a comparative study," *Landslides*, vol. 19, no. 10, pp. 2489–2511, 2022.
- [18] A. N. Chatterjee and B. Ahmad, "A fractional-order differential equation model of COVID-19 infection of epithelial cells," *Solitons & Fractals*, vol. 147, p. 110952, 2021.
- [19] Y. Huang, H. Wu, and E. P. Acosta, "Hierarchical Bayesian inference for hiv dynamic differential equation models incorporating multiple treatment factors," *Biometrical Journal*, vol. 52, no. 4, pp. 470–486, 2010.
- [20] S. Thirumalai, R. Seshadri, and S. Yuzbasi, "Spectral solutions of fractional differential equations modelling combined drug therapy for HIV infection," *Solitons & Fractals*, vol. 151, p. 111234, 2021.
- [21] J. Leander, J. Almqvist, C. Ahlström, J. Gabrielsson, and M. Jirstrand, "Mixed effects modeling using stochastic differential equations: illustrated by pharmacokinetic data of nicotinic acid in obese Zucker rats," *The AAPS Journal*, vol. 17, no. 3, pp. 586–596, 2015.
- [22] M. Fuentes-Gari, R. Misener, M. C. Georgiadis et al., "Selecting a differential equation cell cycle model for simulating leukemia treatment," *Industrial & Engineering Chemistry Research*, vol. 54, no. 36, pp. 8847–8859, 2015.
- [23] A. G. Hawkes, "Hawkes jump-diffusions and finance: a brief history and review," *The European Journal of Finance*, vol. 28, no. 7, pp. 627–641, 2020.
- [24] J. Derek Tucker, L. Shand, and J. R. Lewis, "Handling missing data in self-exciting point process models," *Spatial Statistics*, vol. 29, pp. 160–176, 2019.
- [25] E. W. Fox, F. P. Schoenberg, and J. S. Gordon, "Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences," *The Annals of Applied Statistics*, vol. 10, no. 3, pp. 1725–1756, 2016.
- [26] L. Lesage, M. Deaconu, A. Lejay, J. A. Meira, G. Nichil, and R. State, "Hawkes processes framework with a gamma density as excitation function: application to natural disasters for insurance," *Methodology and Computing in Applied Probability*, vol. 24, pp. 1–29, 2022.
- [27] Z. Sun, W. Dong, J. Shi, and Z. Huang, "Towards predictive analysis on disease progression: a variational Hawkes process model," *IEEE Journal of Biomedical and Health Informatics, Biomedical and Health Informatics*, vol. 25, no. 11, pp. 4195–4206, 2021.
- [28] A. Abouzeid, O.-C. Granmo, C. Webersik, and M. Goodwin, "Learning automata-based misinformation mitigation via Hawkes processes," *Information Systems Frontiers: A Journal of Research and Innovation*, vol. 23, no. 5, pp. 1169–1188, 2021.
- [29] V. Avina-Corral, J. De Jesus Rangel-Magdaleno, H. Peregrina-Barreto, and J. Ramirez-Cortes, "Bearing fault detection in asd-powered induction machine using modwt and image edge detection," *IEEE Access*, vol. 10, pp. 24181–24193, 2022.
- [30] B. Dai, J. Li, J. Zhou et al., "Application of a modified empirical wavelet transform method in vlf/lf lightning electric field signals," *Remote Sensing*, vol. 14, no. 6, p. 1308, 2022.
- [31] Y. Ge, L. Li, and G. Zhang, "A Fourier transform method for solving backward stochastic differential equations,"

- Methodology and Computing in Applied Probability*, vol. 24, no. 1, pp. 385–412, 2022.
- [32] T. Revathi, T. Rajalaxmi, R. Sundara Rajan, and W. P. Freire, “Deep quaternion Fourier transform for salient object detection,” *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 6, pp. 11331–11340, 2021.
- [33] A. Graps, “An introduction to wavelets,” *IEEE Computational Science and Engineering*, vol. 2, no. 2, pp. 50–61, 1995.
- [34] X. Zhao, S. Barber, C. C. Taylor, and Z. Milan, “Classification tree methods for panel data using wavelet-transformed time series,” *Computational Statistics and Data Analysis*, vol. 127, no. 11, pp. 204–216, 2018.
- [35] X. Zhao, S. Barber, C. C. Taylor, and Z. Milan, “Interval forecasts based on regression trees for streaming data,” *Advances in Data Analysis and Classification*, vol. 15, no. 1, pp. 5–36, 2021.
- [36] S. L. Hyland, M. Faltys, M. Hüser et al., “Early prediction of circulatory failure in the intensive care unit using machine learning,” *Nature Medicine*, vol. 26, no. 3, pp. 364–373, 2020.
- [37] A. E. Johnson, T. J. Pollard, L. Shen et al., “Mimic-iii, a freely accessible critical care database,” *Scientific Data*, vol. 3, no. 1, p. 160035, 2016.