

Article

GENAVOS: A New Tool for Modelling and Analyzing Cancer Gene Regulatory Networks Using Delayed Nonlinear Variable Order Fractional System

Hanif Yaghoobi¹, Keivan Maghooli^{1,*}, Masoud Asadi-Khiavi² and Nader Jafarnia Dabanloo¹

¹ Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran; Hanif_yaghoobi@Tabrizu.ac.ir (H.Y); Jafarnia@srbiau.ac.ir (N.J.D)

² Department of Pharmacotherapy, School of Pharmacy, Zanjan University of Medical Sciences, Zanjan, Iran; Masadi@zums.ac.ir (M.A-Kh)

* Correspondence: k_maghooli@srbiau.ac.ir (K.M)

Abstract: Complex diseases such as cancer are caused by changes in the Gene Regulatory Networks. Systems that model the complex dynamics of these networks along with adapting to real gene expression data are closer to reality and can help understand the creation and treatment of cancer. In this paper, for the first time, modelling of gene regulatory networks is performed using delayed nonlinear variable order fractional systems in the state space by a new tool called GENAVOS. This tool uses gene expression time series data to identify and optimize system parameters. This software has several tools for analysing system dynamics. The results show that the nonlinear variable order fractional systems have very good flexibility in adapting to real data. We found that regulatory networks in cancer cells actually have a larger delay parameter than in normal cells. It is also possible to create chaos, periodic and quasi-periodic oscillations by changing the delay, degradation and synthesis rates. Our findings indicate a profound effect of time-varying order on these networks, which may be related to a type of cellular memory due to epigenetic and environmental factors. We showed that by changing the delay parameter and the variable order function for a normal cell system, its behaviour changes and becomes quite similar to the behaviour of a cancer cell. This work also confirms the effective role of the miR-17-92 cluster in the cancer cell cycle. GENAVOS is available at <https://github.com/hanif-y/GENAVOS> with its user guide and MATLAB codes.

Keywords: Gene Regulatory Networks; Non-Linear Variable Order Fractional System; Gene Expression; Epigenetic Memory

1. Introduction

The Gene Regulatory Network (GRN) is a set of molecular regulators that interact with each other and with other cellular materials to control the expression level of mRNA and proteins. The regulator can be DNA, RNA, proteins and their complexes. This interference can be direct or indirect (via transcribed RNA or translated protein) [1]. Recently, the concept of genes has been developed into two categories: protein-encoding genes and non-encoding genes. Micro-RNAs (miRNA) and Long Non-Coding RNAs (LNC-RNA) are non-coding genes. By this latter definition, genes are regulated by other genes, proteins, metabolites, and DNA as well as epigenetic factors in a gene regulatory network [2]. Although the interaction between coding genes is through proteins or metabolites or their combinations, here GRN refers to the network of interactions between genes without considering the intermediates between them [3]. This is a logical way to describe phenomena observed with transcription profiles, such as those occurring with the popular microarray technology or Next generation sequencing (NGS) [3]. Cancer is actually an abnormal process of alteration in the cell molecular regulation network [4]. Therefore, modelling GRNs is one of the important goals of system biology to study the creation, development and treatment of cancer.

In order to model and infer gene networks, gene expression time course data is usually used. These time series reflect intracellular dynamic processes such as time-dependent dynamic of the cell cycle in cancer disease or its treatment [5]. The first step in modelling is to determine the structure of the network. Then, a complete analysis of the dynamics of its components requires solving the Master Equations. By writing the master equations, this network is described as a dynamic system with nonlinear differential equations in state space. In systems theory, GRNs are complex nonlinear dynamic systems with time delays, and complex nonlinear dynamics can be found even within the simplified and constructive infrastructure of these systems [6]. Transcription and translation are processes that take a certain amount of time to complete. Therefore, these two gene expression processes cause time delays in biochemical systems [7]. Delays are inherent in GRN systems and can lead to complex behaviours [6] [8]. The nonlinearity of biochemical phenomena such as phosphorylation and dimerization, auto-regulation or feedback loops are known in GRNs [9]. Also, experimental studies have shown that the function of transcription have a sigmoidal form [10][11]. Environmental factors such as drugs and chemicals, temperature, light and circadian rhythm as well as epigenetic factors such as DNA methylation affect GRNs [12]. Epigenetic mechanism by chromatin modification may provide cellular memory by blocking or allowing transcription [13].

The state variable in GRN system can be concentration of time-varying mRNA (or gene expression) in a biological process such as a cell cycle [5]. Integer-order differential equations do not provide an accurate description of the biological processes of GRNs. Instead, fractional differentials can reflect the memory properties of biochemical phenomena. It has also been shown that transcription dynamics are slow and fractional order systems are very suitable for modelling this dynamics [14]. The fractional order system also provides a large degree of freedom, thus dealing with the nonlinear effects of GRNs. On the other hand, the dynamics of GRNs depend on the fractional order of state space equations. Although some optimal dynamic behaviours are obtained by selecting the appropriate fractional order, due to the high complexity of GRNs, there is a need for more flexible dynamics than fractional order equations [15]. This highly flexible dynamic can be obtained by generalizing fractional order systems to variable order fractional systems (VOFS). In the VOF-GRN system, the fractional order of the differential equations can depend on time, environmental and epigenetic factors. Variable order fractional integrals and derivatives can be operators to describe a system with time-varying memory [16]. These systems are more compatible with biological phenomena [17]. All of the above suggests that VOF systems provide a better description of the dynamics of GRNs.

Yoko Suzuki et al. Showed that time delays completely alter time-dependent dynamics even for the simplest possible circuits with one or two gene elements by self-regulation and regulation of each other. These elements can cause complex behaviours such as periodic, quasi-periodic, weak chaotic, strong and intermittent chaotic dynamics. They introduce a special power spectrum based method for describing and distinguishing these dynamic states. They argued that cancer could cause further delays in the regulation of circadian genes, thus creating a non-periodic dynamic that disrupts the normal process. In fact, the circadian gene regulator circuit can potentially induce chaos in the circuit by delay [6].

Dandan Yue et al investigated a fractional order genetic regulatory network and several of its dynamic behaviours. Their numerical simulations showed the effect of fractional order on stability and oscillations. They presented the results of fraction stability by considering degradation rates as a bifurcation parameter to study the dynamics of the system. Their findings confirm that fractional differential and diffusion can better describe gene activities and provide a greater understanding of nonlinear features [15].

Binbin Tao et al presented a new two-gene regulatory model with a delay parameter using the fractional order system that can better describe the memory and intrinsic properties of genetic regulatory networks. Total delay was selected as the bifurcation parameter and sufficient stability conditions were obtained. In this research, it has been discovered that the network stability interval is inversely proportional to the network fraction order and increasing the order can reduce the

critical amount of delay [18]. A very similar work with similar results is also presented by Qingshan Sun et al [19].

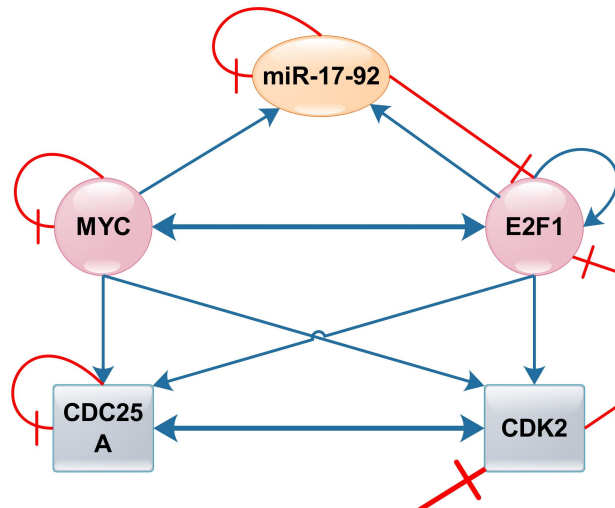


Figure 1. Gene network in the mammalian cancer cell during G1/s transition stage of cell cycle.

In this paper, for the first time, GRNs are modelled using nonlinear delayed VOF systems via gene expression time series data. Also, the dynamic of a real GRN system that fits on real data are discussed. For this purpose, we have created a new tool called GENAVOS (abbreviated Gene Network Analysis by Variable Order Systems). In the first step, time series data is entered into the tool. Then by defining the network structure graph, the nonlinear VOF differential equations of the network are generated in the state space automatically. An advantage of this tool is the ability to approximate the variable-order functions of the system over a period of time. This is possible by selecting several methods as function approximation with free parameters, including the use of Radial Base neural networks, Fourier Series, Constant function, Cosine function and Random function. It is also possible to model networks in which one or more modules do not have time series data. This is done by approximating the expression of dateless genes with the RBF neural network. In the next step, the numerical solution of the VOF system is performed simultaneously with determining and optimizing the system parameters via minimizing Sum of Absolute Error [20]. The parameters identification & optimization is done by the Imperialist Competitive Algorithm (ICA), which is an evolutionary optimization algorithm [21]. After implementing the GRN system, GENAVOS offers several tools for analysing the system dynamics. In the continuation of this article, we present the working methods and the results of this tool.

2. Materials and Methods

2.1. Time Series Dataset

Gene expression time series data of cell cycle process are suitable for analysing GRNs. We used two sets of data. First is related to Human Primary Skin Fibroblasts throughout the Cell Cycle, which is a normal cell [22]. This dataset is available under the access number E-TABM-263 at <https://www.omicsdi.org/dataset>.

The second is the Human Cell Cycle in the HeLa cell line, which is a cancer cell [23]. This dataset is available at <http://genome-www.stanford.edu/Human-CellCycle/HeLa/>.

2.2. Case Study and Network Structure

Because the selected time series data is relevant to the cell cycle, we focus on one of the infrastructures of gene regulatory networks in the cancer cell cycle. In this section, we consider the structure of the cancer gene network regulated by miR-17-92 cluster during G1/s transition in the

mammalian cell cycle [24]. This network is shown in Figure 1. In this graph, T-edges indicate gene inhibition and arrow edges indicate gene activation. This graph is equivalent to adjacency matrix $[a_{ij}]$ where $a_{ij}=0$ means no connection between node i and j , $a_{ij}=1$ means node i is activated by node j and $a_{ij}=-1$ means node i is inhibited by node j . In GENAVOS software, adjacency matrix of network graph must be defined.

2.3. GRN modelling by VOF Systems

The variable fractional order derivative is a generalization of the fractional order derivative. Like the fractional order derivative, several definitions are proposed for the fractional variable order derivative, with Riemann-Liouville, Caputo, and Grünwald–Letnikov being the most common definitions. In this article we have used the left fractional VO derivative of Grünwald–Letnikov. According to this definition we have [25]:

$${}^G D_t^{\alpha(t)} = \lim_{h \rightarrow 0^+} h^{-\alpha(t)} \sum_{j=0}^{\left\lfloor \frac{t}{h} \right\rfloor} (-1)^j \binom{\alpha(jh)}{j} f(t-jh) \quad , \alpha(t) > 0, t \geq 0 \quad (1)$$

$$\binom{\alpha(t)}{j} = \frac{\Gamma(\alpha(t)+1)}{j! \Gamma(\alpha(t)-j+1)} \quad (2)$$

To discretize Equation (1), we assume $t_i = ih$, $i=1,2,\dots,k$, so the numerical approximation of the variable order derivative will be as follows.

$$D_{t_k}^{\alpha(t)} f(t) \approx h^{-\alpha(t_k)} \sum_{j=0}^k (-1)^j \binom{\alpha(t_j)}{j} f(t_{k-j}) \quad (3)$$

Now if a variable order fractional differential equation is as follows [26] ,

$$D_t^{\alpha(t)} y(t) = f(y(t), t) \quad (4)$$

to solve it numerically, we can use Equation 5.

$$y(t_k) = f(y(t_k), t_k) h^{\alpha(t_k)} - \sum_{j=1}^k (-1)^j \binom{\alpha(t_j)}{j} y(t_{k-j}) \quad (5)$$

To model GRN in state space with VOF system, we can write the following equation for each node of the network (genes).

$$D_t^{\alpha(t)} G_i(t) = \sum_{\forall j, a_{ij}=1} \kappa_j F_+(G_j(t-\tau), \theta_j) + \sum_{\forall j, a_{ij}=-1} \kappa_j F_-(G_j(t-\tau), \theta_j) - \gamma_i G_i(t) \quad (6)$$

where F_+ and F_- are nonlinear functions that indicate activation or inhibition of the G_i gene by its control genes, respectively. θ is the parameter of nonlinear functions F , γ is the degradation parameter, κ is the synthesis rate parameter and τ is the delay parameter. The j index represents the genes that activate or inhibit the G_i gene and is determined from the adjacency matrix $[a_{ij}]$ of the network graph [27]. Although F functions can be any nonlinear function, experimental studies have shown that they have a sigmoidal form. The two functions known as the sigmoidal form are the Hill Type and Log-Sigmoid functions, which can be selected in GENAVOS software [28]. Therefore, the functions F_+ and F_- can be in the following forms.

$$F_+(x) = \frac{x^N}{x^N + \theta^N} \quad \text{or} \quad F_+(x) = \frac{1}{1 + e^{-(Nx - \theta)}} \quad (7)$$

$$F_-(x) = \frac{\theta^N}{x^N + \theta^N} \quad \text{or} \quad F_-(x) = \frac{1}{1 + e^{(Nx - \theta)}} \quad (8)$$

Now, to numerically solve Equation 6 using Equation 5 by considering the delay parameter, we first consider the following conditions:

$$\begin{aligned}\tau &= mh, \quad m \in \{0, 1, 2, 3, \dots\} \quad \text{Delay Parameter} \\ T_{sim} &= nh, \quad n \in \{1, 2, 3, \dots\} \quad \text{Simulation Time} \\ G_i(t_l) &= G_i(0), \quad l = -m, -m+1, \dots, 0 \\ f_i(G_i(t-\tau), \kappa, \theta) &= \sum_{\forall j, a_{ij}=1} \kappa_j F_+(G_j(t-\tau), \theta_j) + \sum_{\forall j, a_{ij}=-1} \kappa_j F_-(G_j(t-\tau), \theta_j)\end{aligned}\quad (9)$$

Therefore, the final equation for numerical solution (discrete form) will be as follows:

$$G_i(t_k) = (f_i(G_i(t_{k-m}), \kappa, \theta) - \gamma_i G_i(t_{k-1}))h^{\alpha(t_k)} - \sum_{j=1}^k (-1)^j \binom{\alpha(t_j)}{j} G_i(t_{k-j}) \quad 1 \leq k \leq n \quad (10)$$

2.4. Function Approximation for Variable Order

To solve Equation 6, it is necessary to determine the function of the variable order, $\alpha(t)$. In real systems, $\alpha(t)$ is a time-varying variable (parameter) that affects the system. In the GRN system, for example, this is most likely a time-varying temperature function. For phenomena such as biological systems, it is almost impossible to determine $\alpha(t)$ analytically. So in the optimistic case, $\alpha(t)$ is obtained through measurements as a time series of measurements. However, if the values of $\alpha(t)$ are measured or sampled with the time step of solving equation (h), Equation 6 can be solved numerically and not analytically. But in fact the measurement intervals of an experiment are greater than the time step h . So we have to approximate $\alpha(t)$ from the measured values (time series). GENAVOS has several solutions for determining $\alpha(t)$, including determining $\alpha(t)$ as a constant function, a function with random values, or approximating it by the Fourier series, Radial Base Functions (RBF Neural Network), and the Cosine function. The best results were obtained with RBF neural networks. As we all know, RBF neural networks can approximate any function with acceptable error [29]. Therefore $\alpha(t)$ is approximated as follows:

$$\beta(t) = \sum_{i=1}^L \rho_i e^{-|t-c_i|^2}, \quad \hat{\alpha}(t) = \frac{\beta(t) - \min\{\beta(t) \text{ for all values of } t\}}{\max\{\beta(t)\} - \min\{\beta(t)\} \text{ for all values of } t} \quad (11)$$

where c is the function parameter and the centre of the Gaussian kernel, ρ is the coefficient of the Gaussian kernels and L represents the number of Gaussian functions. The second part of Equation 11 guarantees that $0 \leq \alpha(t) \leq 1$.

2.5. Expression Approximations for Genes that Have No Data

Gene-level regulatory elements include non-coding and coding genes, as seen in Figure 1, which contains the miR-17-92 cluster. Unfortunately, time series data similar to those for coding genes are not available for miRNAs and LNC-RNAs. Here we have included an initiative in GENAVOS to approximate the expression of genes that lack data. The procedure is very similar to the order function approximation described in Section 2.4. Here again, an RBF neural network is considered for the expression of the no-data gene (here miR-17-92) according to Part I of Equation 11. We show the parameters of this network with ζ and μ . So we have:

$$\hat{D}_{non-coding}(t) = \sum_{i=1}^L \zeta_i e^{-|t-\mu_i|^2} \quad (12)$$

2.6. Parameters Identification and Optimization

The purpose of the parameter estimation step is to determine and optimize all system parameters so that the system's solved response (here predicted values of gene expression) has the least error with real data (here gene expression time series data). In GENAVOS this is done by the Imperialist Competitive Algorithm (ICA). ICA is a method in the field of evolutionary computing that finds the optimal answers to various optimization problems. In terms of application, this algorithm falls into the category of evolutionary optimization algorithms such as genetic algorithms, particle swarm optimization method, ant colony algorithm and so on. Like all similar algorithms, the ICA forms the initial set of possible answers. These initial answers are known as "country". ICA, using the Assimilation, Revolution, and Imperialistic Competition operators, gradually improves these initial answers (countries) and finally provides the appropriate answer to the optimization problem [21]. To use this algorithm, the sum of absolute error (SAE) function is defined as a cost function as follows:

$$SAE = \sum_{i=1}^{N_g} \sum_{k=1}^{N_s} \left| \hat{G}_i(t = k \frac{T_s}{h}) - G_i(kT_s) \right| \quad (13)$$

where N_s is the number of time samples of gene expression data, N_g is the number of genes in the network and T_s is the time step of data sampling. The parameters we are trying to optimize are: $\{\kappa, \gamma, \theta, c, \rho, \mu, \zeta\}$.

2.7. GENAVOS Analytical Tools for System Dynamics

After implementing the system and identifying its parameters, we can study the dynamics of the GRN system using GENAVOS analytical tools. The first is related to sensitivity to initial conditions. Here, by applying a coefficient (for example, 1.01 or 0.99), the initial conditions are slightly changed to determine whether the system error changes in a limited way or not. In other words, is the system relatively stable under these initial conditions? Note here that although obtaining the equilibrium point of the system described by Equation 6 is straightforward (obtained by equating the other side of the equation to zero), but examining the stability of the system using the Jacobin matrix characteristic equation is not straightforward due to the existence of a time-varying order. As we know, due to the delay parameter, we need to take the Laplace transform to obtain the characteristic equation of the linearized system [18]. For the fractional order derivative, the Laplace transform is definable as well as the integer order. But for the variable order derivative, the definition of Laplace transform is only possible if $\alpha(t)$ is explicitly expressed as a function [30]. The second is to draw the bifurcation diagram for changing system parameters. These graphs can show which state variables (here genes) change with parameter change and how the system exhibits periodic, quasi-periodic, chaotic, and other behaviours. This diagram can show which parameter change derails the GRN system and leads to disease, for example by changing the delay, degradation, or synthesis rate parameter. It is also possible to check the system's temporal response to changing parameters.

The third is the chaos test with the 0-1 algorithm. This test is used to distinguish regular state from chaotic state in dynamic systems. Unlike the Lyapunov Maximum Exponent test, which uses phase space reconstruction, the 0-1 test uses a time series of data obtained from a dynamic system [31]. The test input is a one-dimensional time series $\varphi(n)$ for $n = 1, 2, \dots$. We use $\varphi(n)$ data to construct the following 2D system:

$$\begin{aligned} p(n+1) &= p(n) + \varphi(n) \cos(cn) \\ q(n+1) &= q(n) + \varphi(n) \sin(cn) \\ n &= 1, 2, \dots, N \end{aligned} \quad (14)$$

where $c \in (0, 2\pi)$ is a random number. The bounded trajectory on the q-p plane indicates a regular system, and the trajectory, similar to Brownian motion or random walking, indicates a chaotic system. The (time-averaged) mean square displacement is defined as follows:

$$M(n) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \left((p(j+n) - p(j))^2 + (q(j+n) - q(j))^2 \right) \quad (15)$$

and its growth rate is as follows:

$$k = \lim_{n \rightarrow \infty} \frac{\log M(n)}{\log n} \quad (16)$$

k is the test result that can be determined by two methods of regression and correlation coefficient [31]. In this work, we have used the correlation coefficient method. According to this method, if k is close to zero, it indicates a regular system, while if it is close to one, it indicates a chaotic system.

3. Results

This section summarizes the results. These results are based on non-cancer cell data to simulate how a normal cell system can turn into cancer. Similar results are given for cancer cells in Supplementary Results.

3.1. Initial Settings and Parameters Tuning

Table 1 shows the initial parameters and initial settings of the system in accordance with what we have already described. ICA settings are listed in Supplementary Table1.

3.2. Numerical Solution and Parameters Identification

Table.2 shows the value of the SAE function in the two modes $\alpha(t)=1$ and $\alpha(t)$ approximated by RBF networks. As can be seen, the SAE differs significantly between the two modes. Table 3 shows the effect of the delay parameter on SAE, so the best results are obtained for $\tau=0.1$ and $\tau=0.3$ for normal and cancer cells, respectively. Figure 2 and Figure 3 show the results of numerical solution of the normal and cancer cell systems with real data points at 2 hour intervals. As shown in Figure 2, 3 the modelled signals fit the data very well and mimic the behaviour of the data. Figure 4 shows the order function curve for each system state variable. Table 4 shows the values of the obtained parameters for the synthesis rate and degradation. There is a similar table for Hela cells in Supplementary Table 2.

Table 1. Initial settings and initial parameters

F_+, F_-	$\alpha(t)$	$Tsim$	T_s	L	h	N_s	N_g	ICA Iteration
<i>Log-sigmoid</i>	RBF-NN	33 Hour	2 Hour	100	0.1	17	8	1000

Table 2. Values of the SAE for the two modes of $\alpha(t)$

<i>Dataset</i>	Hela Cell			Primary Skin Fibroblasts	
$\alpha(t)$	Constant	RBF NN		Constant	RBF NN
SAE	27.82	4.11		18.43	4.94

Table 3. Values of the SAE for the different delay parameter

<i>Dataset</i>	Hela Cell				Primary Skin Fibroblasts			
τ	0.1	0.2	0.3	0.4	0.1	0.2	0.3	0.4
SAE	7.33	6.53	4.11	6.71	4.94	6.68	7.21	7.94

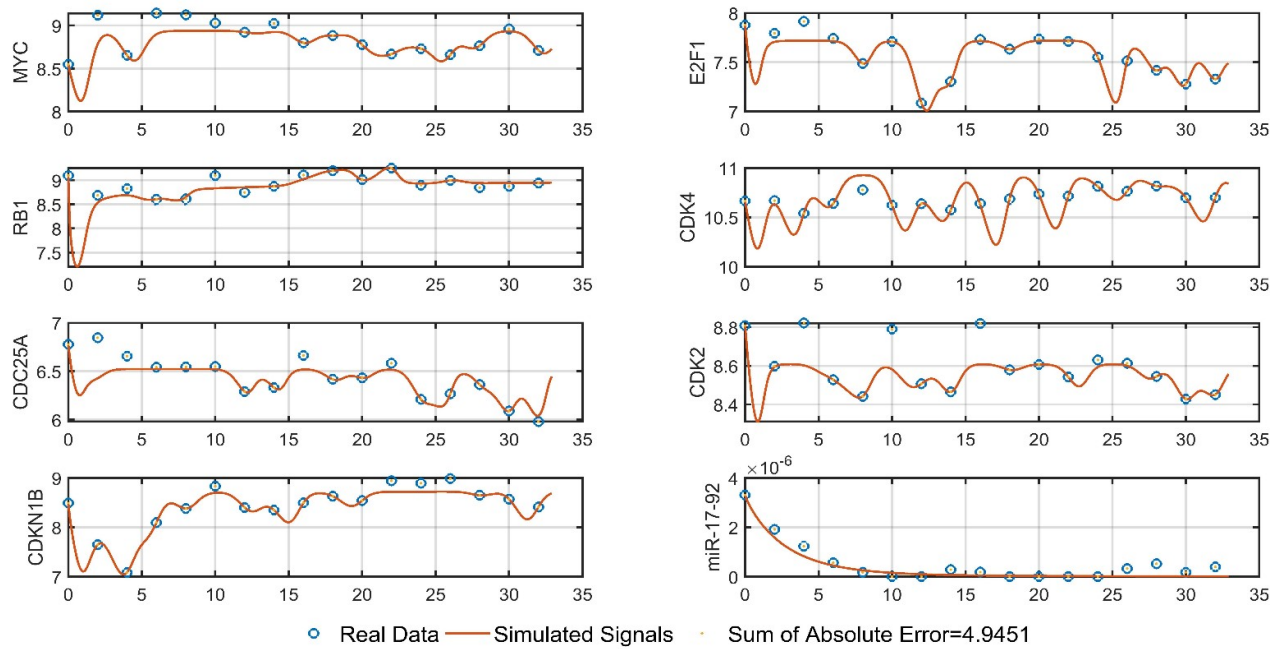


Figure 2. Modeled gene expression signals for Primary Skin Fibroblasts cell in cell cycle. Real data points are marked with a circle.

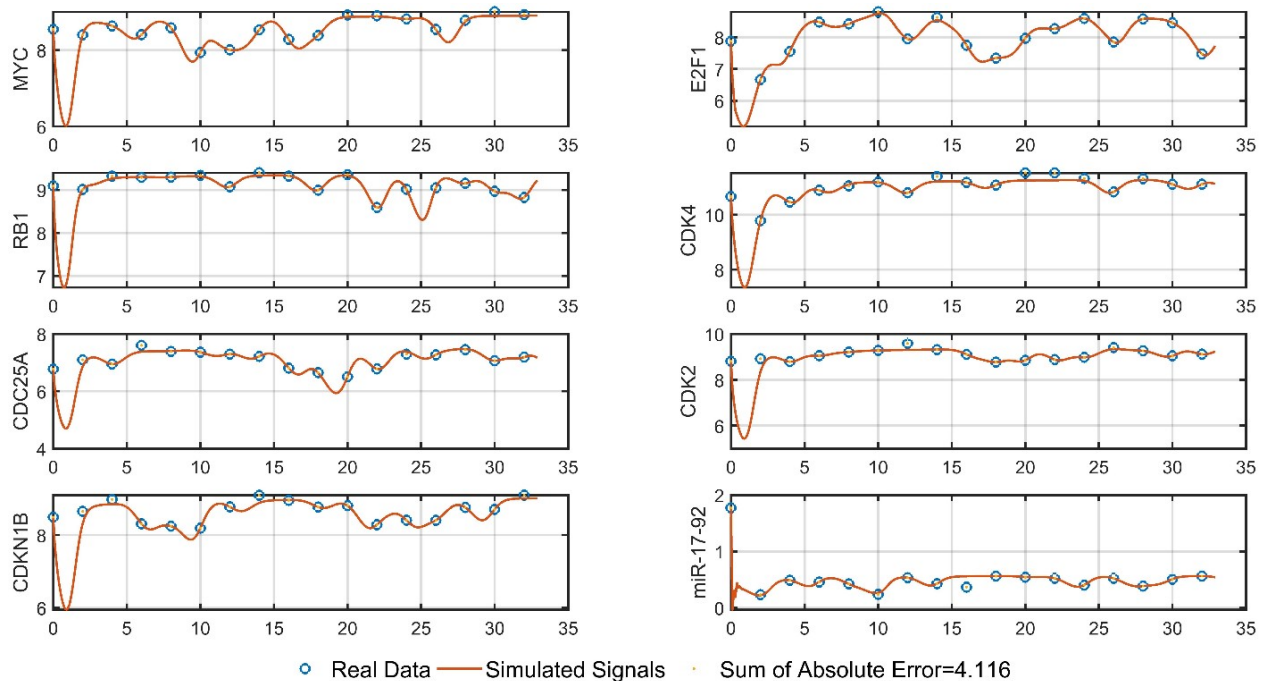


Figure 3. Modeled gene expression signals for HeLa cell in cell cycle. Real data points are marked with a circle.

3.3. System Dynamics Evaluation

Supplementary Figure 4 shows the system response to changing initial conditions. In this figure, the initial conditions are up and down with a coefficient of 1.1 and 0.9. Time delays can cause oscillatory gene expression and provide insights into the dynamics of interactions between genes associated with cancer suppression [4]. Therefore, the delay parameter was selected as a bifurcation diagram parameter. In Table 4, the degradation rate for miR-17-92 is zero.

Because the role of miR-17-92 in the cancer cell cycle is very effective, we chose the degradation rate of miR-17-92 as one of the bifurcation parameters. It has also been shown that a gene with self-inhibiting feedback causes oscillations in the system. Therefore, we chose the MYC synthesis rate as another parameter for the bifurcation diagram. Figure 5 shows the system bifurcation diagram for changing several parameters. The system exhibits complex dynamics similar to chaos or complex oscillations for ranges of parameters. Figure 5(A) and 5(B) shows the behaviour change of the two system state variables (CDC25A, E2F1) for the delay parameter change. Figure 5(E) shows the evolution of the MYC gene to change the Self-Synthesis Rate parameter. Figure 5(F) shows the change in behaviour of E2F1 to change the miR-17-92 degradation rate. Figure 5(C) and 5(D) shows the effect of the variable order $\alpha(t)$. To achieve this goal, we have defined the K parameter as Equation 17:

$$\alpha_{new}(t) = \alpha_{Normal}(t) - K(\alpha_{Cancer}(t) - \alpha_{Normal}(t)) \quad (17)$$

where $\alpha_{Normal}(t)$ is obtained from modelling normal cell data (Primary Skin Fibroblasts) and $\alpha_{Cancer}(t)$ is obtained from modelling cancer cell data (Hela Cell) by the proposed method. Figure 6 shows the modelled time signals corresponding to the bifurcation diagram of Figure 5.

Finally, to check for chaotic dynamics in the model, we used the 0-1 test [31]. The q-p plots of normal cell system are shown in Figure 7. Table 5 shows the results of this test in different system conditions. The q-p plots corresponding to Table 5 are shown in Figure 8.

4. Discussion

Table 2 shows that $\alpha(t)$ has created more flexible dynamics in the time-varying mode. It seems very difficult or impossible to fit the system response on the data without a variable-order differential, while Figures 2 and 3 show that the VOF system fits well on the data. Of course, it is clear that the data may also have noise values, however, several articles have explicitly emphasized the existence of oscillating dynamics and periodic behaviours of gene expression data in the cell cycle, and this has been biologically proven [22][23], [32]. Thus, fluctuations in gene expression over the course of the cell cycle are largely related to the intrinsic properties of the system. Also, Mahboobeh Ghorbani et al have suggested that gene expression is not random, because the Hurst exponent for most of the genes is significantly higher than 0.5. So the genes time series cannot be regarded as a random process [33]. An important point about Figures 2 and 3 is about miR-17-92. As shown in Figure 2, this module is under-expressed in normal cells. Instead, it can be seen in Figure 3 that its expression has increased significantly and has fluctuating states during the cell cycle.

Table 4. values of the obtained parameters of normal cell system

Target Genes	Synthesis Rate of Controller Genes								Degradation of Target Genes
	MYC	E2F1	RB1	CDK4	CDC25A	CDK2	CDKN1B	miR-17-92	
MYC	2.46	0.46	--	--	--	--	--	--	0.26
E2F1	1.96	0.25	5	--	--	--	--	4.02	1.15
RB1	--	--	--	2.47	--	3.24	--	--	0.65
CDK4	--	--	--	3.78	--	--	--	--	0.34
CDC25A	4.2	4.64	--	--	5	4.84	--	--	1.4
CDK2	3.91	0.14	--	--	0.51	--	4.98	--	0.63
CDKN1B	--	--	--	--	--	2.87	--	--	0.3
miR-17-92	1.73	4.5	--	--	--	--	--	0	0

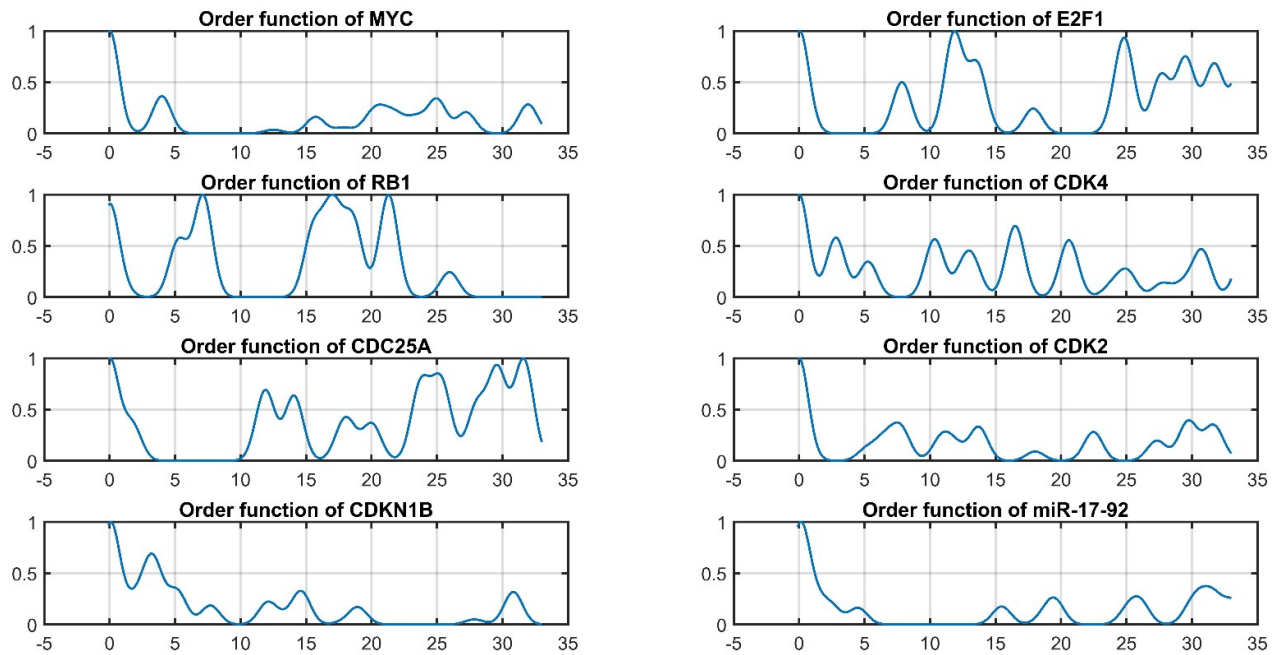


Figure 4. Variable order functions approximated by RBF neural network.

Many studies show that the miR-17-92 cluster is overexpressed in cancer. This is specifically stated in a review article by Xinju Zhang et al [34].

We find two important interpretations of Figure 4. The first is the existence of very slow dynamics in regulatory networks. We calculated the mean of the function $\alpha(t)$ in both normal and cancer cell states. Its maximum was 0.37 and 0.31, respectively, and in general it seems that the dynamics in the cancerous state has slowed down. If we consider $\alpha(t)$ as a system input (and not a system variable [30]) that is affected by epigenetic memory, the difference between these two numbers between cancerous and normal states will be significant [35]. The second is the frequent zeroing of the $\alpha(t)$ function over time, which evokes the type of on-off switching mentioned in the work of Richard P. Halley-Stott et al [36]. It may also support studies that claim that gene expression is inherently discrete [37]. Table 3 confirms that the delay parameter in cancer cell GRNs is larger than in normal cells. This confirms studies that have shown that cancer increases the delay parameter [6].

Table 5. Results of 0-1 test in different conditions of system.[illegible]

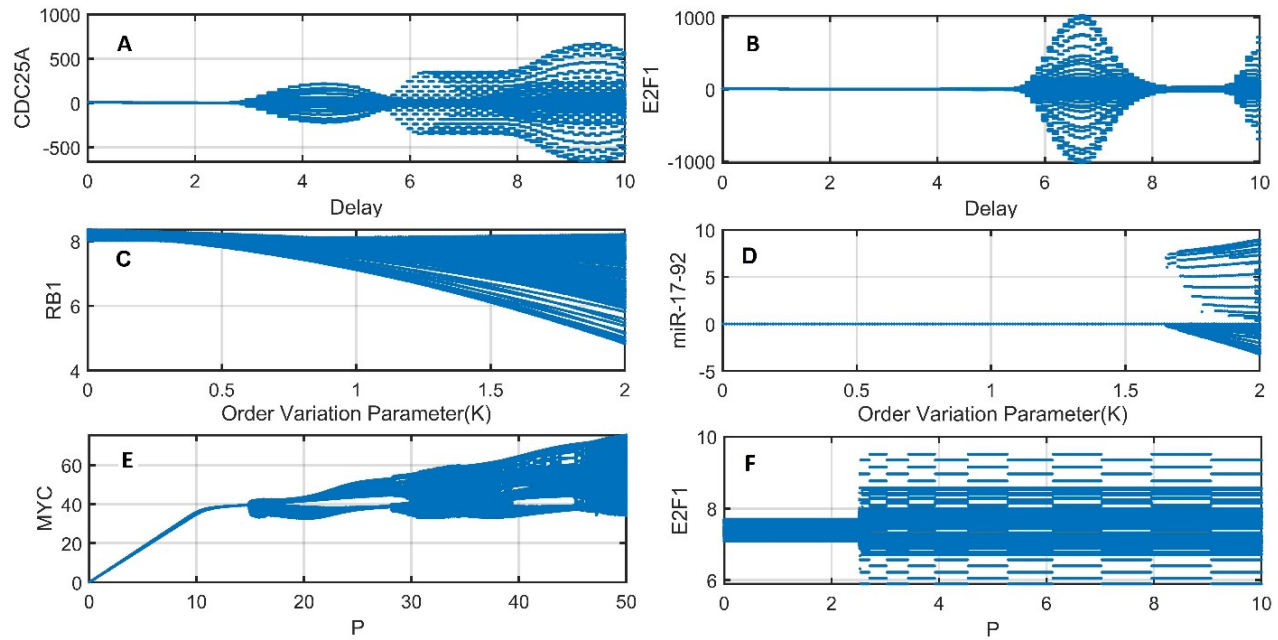


Figure 5. GRN System bifurcation diagram for changing the parameters: (A) , (B) Delay parameter τ . (C), (D) Variable Order parameter K . (E) MYC Self-Synthesis Rate κ_1 . (F) miR-17-92 Degradation parameter γ_8 .

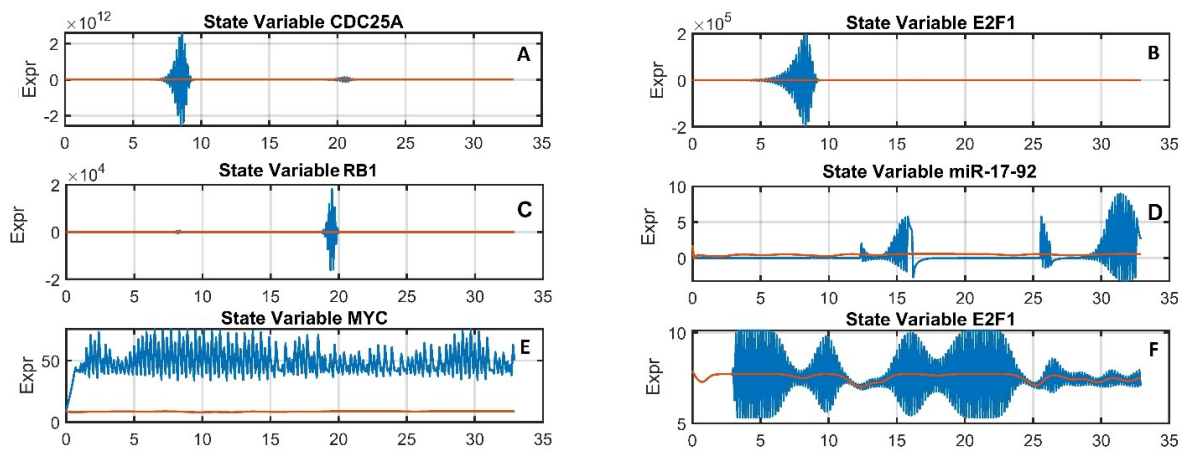


Figure 6. Modelled signals for changed parameters: (A) , (B) Delay parameter $\tau = 10$, (C), (D) Variable Order parameter $K = 2$, (E) MYC Self-Synthesis Rate $\kappa_1 = 50$, (F) miR-17-92 Degradation parameter $\gamma_8 = 10$.

As we expected, Figure 5 shows that the system undergoes a fundamental change by changing its parameters. The E2F1 gene undergoes complex oscillations from the $\tau \approx 5.31$ onwards, while returning to regular behavior between $\tau \approx [8.37, 9.27]$. The oscillation starts again from 9.27. A similar behavior is seen for CDC25A. Oscillation start at around 2.7 and become almost regular at ≈ 5.59 and then oscillate again. A definite opinion as to whether the system of these oscillations is chaotic or quasi-periodic is not straightforward by simply using a bifurcation diagram without analytical methods. That's why we used the 0-1 test. Bifurcation diagram of MYC is similar to the bifurcation diagram of two-dimensional maps (such as Logistic Map). This is due to the presence of a negative feedback loop of the MYC gene on itself, which behaves like a recursive map in the simulation.

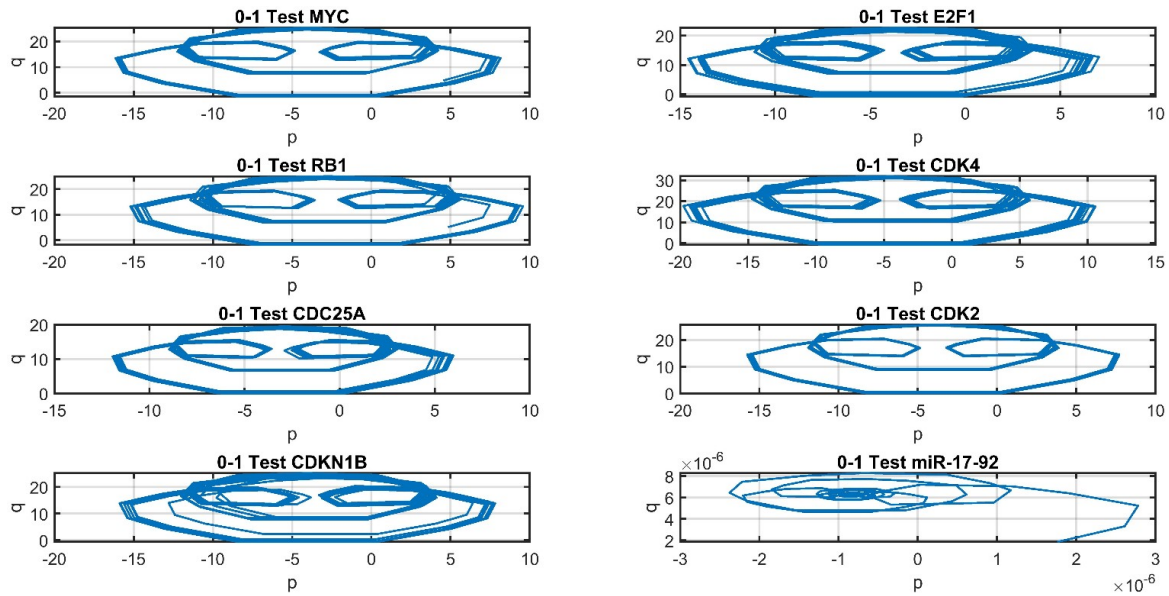


Figure 7. q-p plots of normal cell system without change in parameters

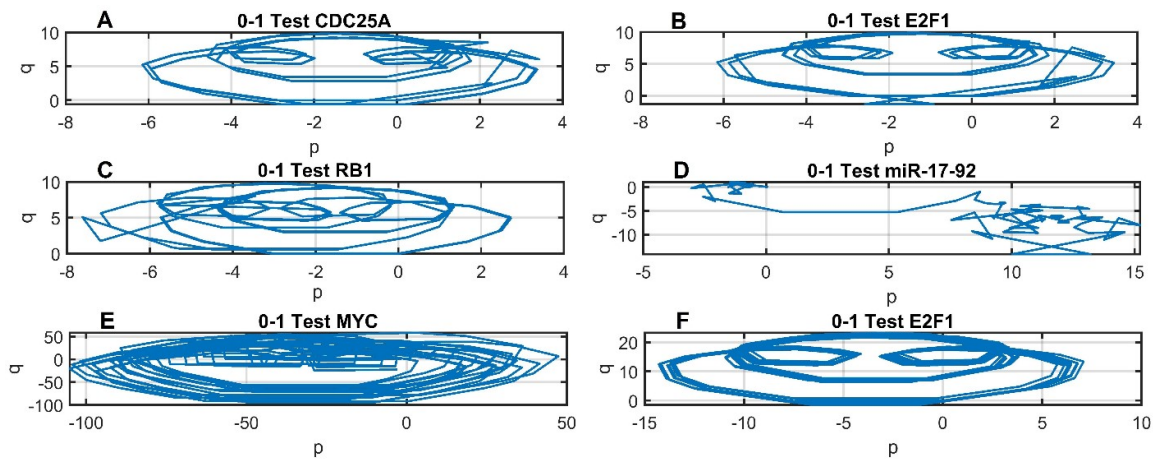


Figure 8. q-p plots corresponding to Table 5

The E2F1 gene oscillates by changing the degradation parameter of miR-17-92 from about 2.51, due to the presence of an inhibitory feedback loop between them. In Figure 6, the time signals confirm the behaviours of the system in a range of parameters where it is irregular corresponding to Figure 5.

A comparison of Figures 7 and 8 clearly shows that changing the parameters corresponding to Figure 6 makes the behaviors of the trajectories irregular and resembling a random state. Table 5 also confirms this. So the signals in Figure 6 are somewhat chaotic. But given their k -values in Table 5, this is probably Weak chaos.

Finally, a question arises: Is it possible to simulate the behavior of a cancer cell by changing some parameters of the normal cell system? To answer this, we first change the delay parameter of the normal cell system from 0.1 to 0.3. Then in Equation 17, we take K equal to 1. When $K = 1$, $\alpha_{new}(t)$ will be equal to $\alpha_{Cancer}(t)$. Figure 9 shows the time response of the system to substituting $\alpha_{Normal}(t)$ with $\alpha_{Cancer}(t)$. We found that the normal cell system shows exactly the same behavior as a cancer cell by replacing $\alpha(t)$ with $\alpha_{Cancer}(t)$. This indicates that the VOF-GRN system is strongly influenced by the variable-order function. The order functions are not parameters that relate to just its own state variables, but system inputs that affect the system as a whole.

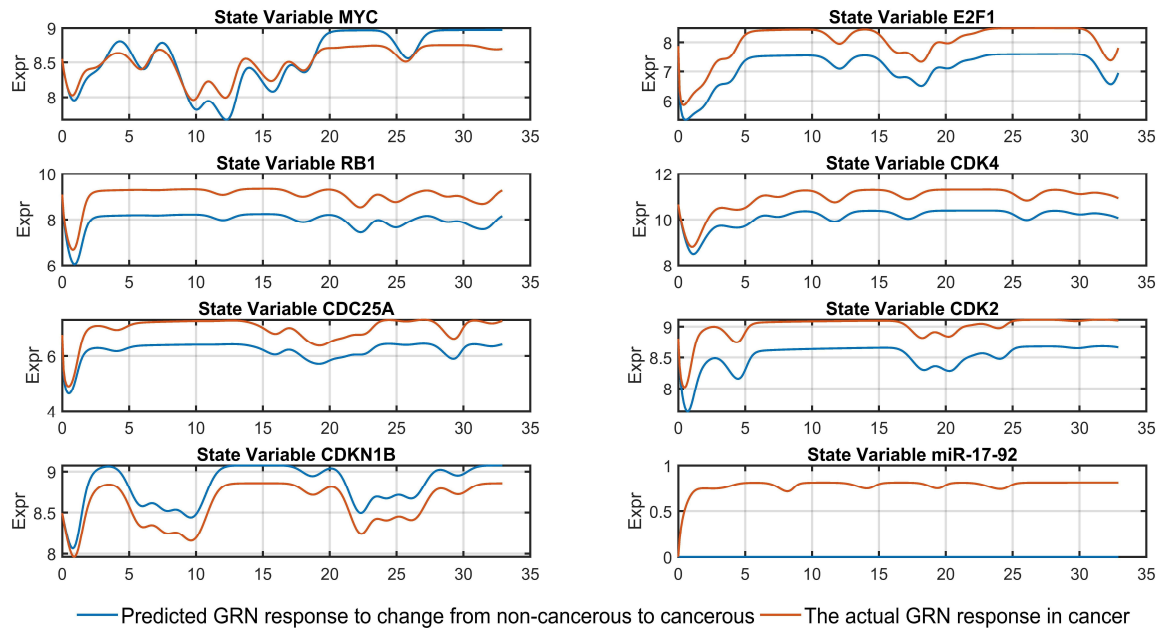


Figure 9. Normal cell shows exactly the same behavior as a cancer cell by replacing $\alpha(t)$ with

$$\alpha_{Cancer}(t)$$

Table 8 also has an important result: k is larger for the cancerous condition than normal. In other words, the weak chaotic property of the system in the cancerous state is more than the normal state.

5. Conclusions

Recently, the theory and application of variable-order differential equations has expanded. These equations are very useful for describing flexible and reality-based systems. In this paper, a new tool called GENAVOS was introduced to model gene regulatory networks. Our results show that the use of delayed nonlinear VOF systems has very flexible dynamics that correspond to real biology. The VOF system for GRN is so strongly influenced by the time-varying order function that it easily goes from normal to cancerous by changing the system's variable order. It is suggested that the variable-order function be more accurately interpreted physically and biologically, and that a solution be provided for its experimental determination.

Supplementary Materials: Genavos Help.

Author Contributions: H.Y. has presented the main idea of this article. Feasibility study and its scientific framework have been done by H.Y., K.M. and N.J.D. Implementation in MATLAB and numerical solution has been done by H.Y. and with the approval and supervision of K.M., N.J.D. M.A-Kh. has worked on the biological part of the paper and matching the findings with biological reality.

Funding: This research received no external funding.

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

References

- [1] D. S. Latchman, "Inhibitory transcription factors," *International Journal of Biochemistry and Cell Biology*, vol. 28, no. 9. Elsevier Ltd, pp. 965–974, 1996, doi: 10.1016/1357-2725(96)00039-8.
- [2] K. Han and J. Lee, "GeneNetFinder2: Improved Inference of Dynamic Gene Regulatory Relations with Multiple Regulators," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 13, no. 1, pp. 4–11, Jan. 2016, doi: 10.1109/TCBB.2015.2450728.
- [3] P. Brazhnik, A. De La Fuente, and P. Mendes, "Gene networks: How to put the function in genomics,"

- Trends in Biotechnology*, vol. 20, no. 11. Trends Biotechnol, pp. 467–472, Nov. 01, 2002, doi: 10.1016/S0167-7799(02)02053-X.
- [4] K. Parmar, K. B. Blyuss, Y. N. Kyrychko, and S. J. Hogan, "Time-Delayed Models of Gene Regulatory Networks," *Computational and Mathematical Methods in Medicine*, vol. 2015. Hindawi Publishing Corporation, 2015, doi: 10.1155/2015/347273.
 - [5] H. Yaghoobi, S. Haghipour, H. Hamzeiy, and M. Asadi-Khiavi, "A review of modeling techniques for genetic regulatory networks," *Journal of Medical Signals and Sensors*, vol. 2, no. 1. Isfahan University of Medical Sciences(IUMS), pp. 61–70, Jan. 01, 2012, doi: 10.4103/2228-7477.108179.
 - [6] Y. Suzuki, M. Lu, E. Ben-Jacob, and J. N. Onuchic, "Periodic, Quasi-periodic and Chaotic Dynamics in Simple Gene Elements with Time Delays," *Sci. Rep.*, vol. 6, no. 1, p. 21037, Aug. 2016, doi: 10.1038/srep21037.
 - [7] S. Nikolov, J. V. Gonzalez, M. Nenov, and O. Wolkenhauer, "Dynamics of a miRNA model with two delays," *Biotechnol. Biotechnol. Equip.*, vol. 26, no. 5, pp. 3315–3320, Oct. 2012, doi: 10.5504/bbeq.2012.0067.
 - [8] Z. Levnajić and B. Tadić, "Stability and chaos in coupled two-dimensional maps on gene regulatory network of bacterium *E. coli*," *Chaos An Interdiscip. J. Nonlinear Sci.*, vol. 20, no. 3, p. 033115, Sep. 2010, doi: 10.1063/1.3474906.
 - [9] D. Yue, Z.-H. Guan, J. Chen, G. Ling, and Y. Wu, "Bifurcations and chaos of a discrete-time model in genetic regulatory networks," *Nonlinear Dyn.*, vol. 87, no. 1, pp. 567–586, Jan. 2017, doi: 10.1007/s11071-016-3061-1.
 - [10] A. Ahmed and E. I. Verriest, "Modeling & Analysis of Gene Expression as a Nonlinear Feedback Problem With State-Dependent Delay," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 12679–12684, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.2248.
 - [11] J. J. Tyson and B. Novak, "Regulation of the eukaryotic cell cycle: Molecular antagonism, hysteresis, and irreversible transitions," *J. Theor. Biol.*, vol. 210, no. 2, pp. 249–263, May 2001, doi: 10.1006/jtbi.2001.2293.
 - [12] "Environmental Influences on Gene Expression | Learn Science at Scitable." <https://www.nature.com/scitable/topicpage/environmental-influences-on-gene-expression-536/> (accessed Sep. 07, 2020).
 - [13] S. Berry and C. Dean, "Environmental perception and epigenetic memory: mechanistic insight through FLC," 2015, doi: 10.1111/tpj.12869.
 - [14] K. Wei, S. Gao, S. Zhong, and H. Ma, "Fractional dynamics of globally slow transcription and its impact on deterministic genetic oscillation," *PLoS One*, vol. 7, no. 6, Jun. 2012, doi: 10.1371/journal.pone.0038383.
 - [15] D. Yue, Z.-H. Guan, M. Chi, B. Hu, Z.-W. Liu, and J. Chen, "Stability and Hopf bifurcation of fractional genetic regulatory networks with diffusion," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 10443–10448, Jul. 2017, doi: 10.1016/J.IFACOL.2017.08.1973.
 - [16] H. Sheng, H. G. Sun, C. Coopmans, Y. Q. Chen, and G. W. Bohannan, "A Physical experimental study of variable-order fractional integrator and differentiator," *Eur. Phys. J. Spec. Top. 2011 1931*, vol. 193, no. 1, pp. 93–104, Apr. 2011, doi: 10.1140/EPJST/E2011-01384-4.
 - [17] W. G. Glöckle and T. F. Nonnenmacher, "A fractional calculus approach to self-similar protein dynamics," *Biophys. J.*, vol. 68, no. 1, pp. 46–53, 1995, doi: 10.1016/S0006-3495(95)80157-8.
 - [18] B. Tao, M. Xiao, Q. Sun, and J. Cao, "Hopf bifurcation analysis of a delayed fractional-order genetic regulatory network model," *Neurocomputing*, vol. 275, pp. 677–686, Jan. 2018, doi: 10.1016/J.NEUCOM.2017.09.018.
 - [19] Q. Sun, M. Xiao, and B. Tao, "Local Bifurcation Analysis of a Fractional-Order Dynamic Model of Genetic Regulatory Networks with Delays," *Neural Process. Lett.*, vol. 47, no. 3, pp. 1285–1296, Jun. 2018, doi: 10.1007/s11063-017-9690-7.
 - [20] J. Nowaková and M. Pokorný, "System Identification Using Genetic Algorithms," in *Advances in Intelligent Systems and Computing*, 2014, vol. 303, pp. 413–418, doi: 10.1007/978-3-319-08156-4_41.
 - [21] E. Atashpaz-Gargari and C. Lucas, "Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition," in *2007 IEEE Congress on Evolutionary Computation*, Sep. 2007, pp. 4661–4667, doi: 10.1109/CEC.2007.4425083.
 - [22] Z. Bar-Joseph *et al.*, "Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 3, pp. 955–960, Jan. 2008, doi: 10.1073/pnas.0704723105.
 - [23] M. L. Whitfield *et al.*, "Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors," *Mol. Biol. Cell*, vol. 13, no. 6, pp. 1977–2000, Jun. 2002, Accessed: Aug. 14, 2019. [Online]. Available: <https://www.molbiolcell.org/doi/10.1091/mbc.02-02-0030>.

- [24] L. Yang *et al.*, "Robustness and Backbone Motif of a Cancer Network Regulated by miR-17-92 Cluster during the G1/S Transition," *PLoS One*, vol. 8, no. 3, p. e57009, Mar. 2013, doi: 10.1371/journal.pone.0057009.
- [25] H. Sun, A. Chang, Y. Zhang, and W. Chen, "A review on variable-order fractional differential equations: Mathematical foundations, physical models, numerical methods and applications," *Fractional Calculus and Applied Analysis*, vol. 22, no. 1. De Gruyter, pp. 27–59, Feb. 25, 2019, doi: 10.1515/fca-2019-0003.
- [26] "Fractional-Order Nonlinear Systems - Modeling, Analysis and Simulation | Ivo Petras | Springer." <https://www.springer.com/gp/book/9783642181009> (accessed Jul. 19, 2020).
- [27] J. Cao, X. Qi, and H. Zhao, "Modeling gene regulation networks using ordinary differential equations," *Methods Mol. Biol.*, vol. 802, pp. 185–197, 2012, doi: 10.1007/978-1-61779-400-1_12.
- [28] E. R. Dougherty, *Genomic signal processing and statistics*. Hindawi Pub. Corp, 2005.
- [29] N. Mai-Duy and T. Tran-Cong, "Approximation of function and its derivatives using radial basis function networks," *Appl. Math. Model.*, vol. 27, no. 3, pp. 197–220, Mar. 2003, doi: 10.1016/S0307-904X(02)00101-4.
- [30] S. Sahoo, S. Saha Ray, S. Das, and R. K. Bera, "The formation of dynamic variable order fractional differential equation," *International Journal of Modern Physics C*, vol. 27, no. 7. World Scientific Publishing Co. Pte Ltd, Jul. 01, 2016, doi: 10.1142/S0129183116500741.
- [31] G. A. Gottwald and I. Melbourne, "The 0-1 test for chaos: A review," in *Lecture Notes in Physics*, vol. 915, Springer Verlag, 2016, pp. 221–247.
- [32] G. D. Grant *et al.*, "Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors," *Mol. Biol. Cell*, vol. 24, no. 23, pp. 3634–3650, Dec. 2013, doi: 10.1091/mbc.E13-05-0264.
- [33] M. Ghorbani, E. A. Jonckheere, and P. Bogdan, "Gene Expression Is Not Random: Scaling, Long-Range Cross-Dependence, and Fractal Characteristics of Gene Regulatory Networks," *Front. Physiol.*, vol. 9, no. OCT, p. 1446, Oct. 2018, doi: 10.3389/fphys.2018.01446.
- [34] X. Zhang, Y. Li, P. Qi, and Z. Ma, "Biology of MiR-17-92 cluster and its progress in lung cancer," *Int. J. Med. Sci.*, vol. 15, no. 13, pp. 1443–1448, Sep. 2018, doi: 10.7150/ijms.27341.
- [35] M. Kim and J. Costello, "DNA methylation: An epigenetic mark of cellular memory," *Experimental and Molecular Medicine*, vol. 49, no. 4. Nature Publishing Group, Apr. 07, 2017, doi: 10.1038/emmm.2017.10.
- [36] R. P. Halley-Stott and J. B. Gurdon, "Epigenetic memory in the context of nuclear reprogramming and cancer," *Brief. Funct. Genomics*, vol. 12, no. 3, pp. 164–173, May 2013, doi: 10.1093/bfgp/elt011.
- [37] J. Goutsias and S. Kim, "A Nonlinear Discrete Dynamical Model for Transcriptional Regulation: Construction and Properties," *Biophys. J.*, vol. 86, no. 4, pp. 1922–1945, Apr. 2004, doi: 10.1016/S0006-3495(04)74257-5.