AIMS Mathematics, 11(11): 38850–38894.
DOI: 10.3934/math.20261499
Received: 27 August 2026
Revised: 18 October 2026
Accepted: 06 November 2026
Published: 20 November 2026

http://www.aimspress.com/journal/Math

Research article

A novel quantum neural network approach for analyzing deterministic and stochastic breast cancer cardiotoxicity model

Azizah Alrajhi[1,*], **and Misbah Ullah**[2,*]

[1] Department of Mathematics, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

[2] Department of Mathematics, City University of Science and Information Technology, Peshawar, KP, 25000, Pakistan

* **Correspondence:** Email: misbahullahkhan146@gmail.com, aalrajhi@kau.edu.sa.

Abstract: Breast cancer remains one of major global public health threats, imposing a significant burden on human health and contributing potentially to cancer-related morbidity and mortality worldwide. Quantum neural networks are emerging as promising nonlinear learning frameworks for approximating complex biological dynamic systems, offering potential advantages in representing intricate nonlinear relationships. In the present study, we investigate a hybrid quantum classical neural network for approximating deterministic and stochastic trajectories of a five-compartments breast cancer mathematical model. The quantum component contains five qubits and three temporal data re-uploading layers. Fifteen Pauli expectation features are transferred to a classical network with two hidden layers of 24 neurons and five outputs. In the deterministic regime, the hybrid quantum neural network (HQNN) achieves coefficient of determination values ranging from 0.999367 to 0.999836, normalized mean absolute errors below 0.46%, and normalized root mean square errors below 0.70%. For the selected stochastic realization, the corresponding R^2 values range from 0.986659 to 0.996558, while the normalized mean absolute errors and normalized root-mean-square errors remain below 1.89% and 2.48%, respectively. An additional architecture investigation comprised 48 training runs, while five independently simulated stochastic trajectories were each fitted using three network initializations. Across the five paths, the mean held-out normalized root mean square error was 0.2238, with a 95% confidence interval of $[0.1614, 0.2861]$. The corresponding mean held-out coefficient of determination was -26.64, indicating limited accuracy on the withheld temporal intervals despite the favorable original reconstruction results. The findings illustrate the approximation behavior and current limitations of a simulated hybrid quantum-classical surrogate.

Keywords: breast cancer; cardiotoxicity; stochastic mathematical model; hybrid quantum neural network; temporal validation; numerical approximation
Mathematics Subject Classification: 60H10, 81P68, 65C30

1. Introduction

1.1. Epidemiological and clinical motivation

Breast cancer develops when cells in breast tissue undergo uncontrolled proliferation, leading to tumor formation [1]. According to Global Burden of Disease Cancer Collaboration, the incidences of breast cancer have the highest prevalence rate compared with any other type of cancer [2, 3]. Breast cancer can occur among post-pubescent females, regardless of age or country, but the prevalence rate usually rises with age [4]. The exact causes of breast cancer are not fully known yet; however, various risks have been identified. These include genetics, hormones, environment, and lifestyle habits [5]. The number of female breast cancer diagnoses was recorded to be 2.3 million cases in 2020. Moreover, there were almost 685000 deaths caused by this deadly disease [6]. By the end of 2020, the number of breast cancer survivors among the population was found to be around 7.8 million individuals diagnosed within five years [6]. It means that breast cancer is the most frequently diagnosed cancer worldwide. The burden of disability-adjusted life years (DALYs) attributable to breast cancer was also higher than that of other cancers among women. Significant advances have been made regarding breast cancer research related to the initiation, growth, progression, spreading, diagnosis, and therapy of the disease [7, 8]. However, although significant achievements have been made during the last five decades, this disease continues to pose serious problems to public health around the world. At present, breast cancer ranks first among the most diagnosed invasive cancers among females globally [9]. According to statistical figures published by the World Health Organization (WHO), breast cancer is the most diagnosed cancer type, followed by leukemia as the second cancer-related death cause in the country in 2023 [10] . Breast cancer incidence rates have increased consistently over time in Saudi Arabia compared with those observed in many Western nations. The investigation of the epidemiology of breast cancer, especially concerning age-related and geographic differences, is thus essential in the research agenda of Saudi Arabia. The significance of such research is even more pertinent because a significant number of the women in the world are relatively young. For instance, the epidemiological study by [11] estimated the breast cancer burden across different countries from 2001 to 2008.

1.2. Previous mathematical and computational investigations

Mathematical modeling is a structured tool that allows the representation of complicated biological phenomena, investigation of interactions between variables connected to a certain disease, prediction of system's behavior, and assessment of possible control measures [12–15]. The use of mathematical modeling in cancer studies originates from 1954 [16]. Since then, mathematical models were widely used for studying tumor' initiation, growth, progression, immune response, metastasis, and cancer treatment [17–19]. A mathematical model for analyzing chemotherapy and its influence on the tumor treatment was created in [20]. The models may take interactions between populations of tumor cells, energy or nutrient resources, chemotherapy, and responses to it accounts. More recently, a model with few parameters was designed for the analysis of low-dose chemotherapy and signaling by angiogenesis between blood vessels and malignant tumors. A compartmental model including

populations of malignant and healthy cells was proposed by Jordao and Tavares [21]. A mathematical model of the interaction between tumors and immune system aiming at creating better cancer treatment strategies was designed by Mahlbacher [22]. Many other mathematical models of cancer development, progression, and treatment have been proposed [23–25].

Despite being powerful tools to solve mathematical models for cancer using numerical methods, repeated numeric integration can be highly computational in terms of computing the solutions for different parameterizations, uncertainties, optimization, or therapy scenarios. Due to recent developments in advanced machine learning methods, studies have been conducted on designing data-driven surrogate models to estimate solutions for nonlinear dynamic systems [26–28]. After training, the neural-network surrogates can quickly produce solutions for the trajectory of the systems without repeated integration of the system of equations [29]. Nevertheless, the traditional neural-network may need large modeling capacity to capture relevant information from such time-series data.

In last decades, the development of novel neuro-computing techniques has attracted considerable attention across various real-world applications, including complex epidemiological modeling [30–33]. Recent applications of artificial neural-networks can be found in [34, 35]. The hybrid quantum neural network represents another type of computing framework, where a parameterized quantum circuit serves as a trainable nonlinear feature extractor, and a classical neural-network learns to map the quantum expectation values from the former to the necessary outputs [36, 37]. Using input-dependent rotation, variational gate application, entanglement between qubits, and data re-uploading, parameterized quantum circuits can generate bounded, nonlinear, and correlated feature representations, which may find use in multivariate regression problems and dynamic system approximation [38]. Nevertheless, it needs to be shown with numerical experiments compared with classical methods whether such representation brings any practical benefit. In the current research, a parameterized quantum circuit is analyzed with a classical state vector simulator. Consequently, the usage of an HQNN is driven by its representational capabilities, rather than the computation advantage.

1.3. Purpose and scope of the present investigation

In light of the prior discussion, the present study is about the novel implementation of an end-to-end supervised HQNN surrogate for the numerical trajectories of a cancer cardiotoxicity five-compartment model. The system of ordinary differential equations (ODEs) represents the coupled dynamics of five state variables representing tumor and treatment-induced dynamics, including the cardiotoxicity arising from cancer therapy. Reference solutions for the deterministic and stochastic system are obtained by the Legendre spectral collocation method. In turn, these solutions are used as supervised targets to train the HQNN surrogate model. The suggested HQNN takes normalized time as the classical input and computes its function with a five-qubit quantum circuit with three layers for data re-uploading. In each of the layers, frequency-based temporal encoding that depends on qubits is applied, then trainable RX, RY and RZ rotations are performed and circular entanglement with the controlled NOT (CNOT) gates is established. The measurements of the Pauli-X, Pauli-Y, and Pauli-Z observable of all the five qubits result in 15 quantum expectation features. The latter are fed into the classical regression head with a 15–24–24–5 architecture, which estimates the trajectories of the five compartments of the model.

The key contributions of this study are summarized as follows:

- A five-compartment mathematical framework is used to describe the coupled dynamics of cancer's progression, therapeutic response, and treatment-associated cardiotoxicity.
- Deterministic and stochastic reference trajectories are generated using the Legendre spectral collocation method (LSCM).
- A five-qubit HQNN is developed as a multi-output surrogate for simultaneously approximating the trajectories of all five model compartments. Its quantum component incorporates frequency-based temporal encoding, three data re-uploading layers, trainable single-qubit rotations, circular CNOT entanglement, and 15 Pauli expectation features.
- The quantum feature extractor and the classical 15–24–24–5 regression head are trained jointly within an end-to-end supervised optimization framework.
- The approximation performance of the HQNN is evaluated over the complete temporal grid using the mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), normalized error measures, and coefficient of determination (R^2).

2. Model formulation

We consider an existing five-compartment mathematical model for breast cancer's progression, recovery following chemotherapy, disease recurrence and treatment-associated cardiotoxicity. The patient population is partitioned into compartments $X(t)$, $B(t)$, $C(t)$, $R(t)$, and $E(t)$, based on the status of the disease and the impact of the treatment. In particular, $X(t)$ denotes the patients suffering from Stage I and Stage II breast cancer, while $B(t)$ and $C(t)$ denote patients suffering from Stage III and Stage IV cancer, respectively. The compartment $R(t)$ includes the patients who recover from the disease after undergoing chemotherapy, whereas $E(t)$ includes the patients suffering from chemotherapy-induced cardiotoxicity. The compartmental model is based on the clinical scenarios outlined in [39].

The corresponding deterministic system is given by

$$\begin{cases} \dfrac{dX}{dt} = \Lambda_1 - (\phi_1 + \beta_2)X, \\ \dfrac{dB}{dt} = \Lambda_2 + \beta_2 X + \psi_1 R - (\phi_2 + \beta_1 + \tau_1 + \mu_1)B, \\ \dfrac{dC}{dt} = \Lambda_3 + \beta_1 B + \psi_2 R - (\phi_3 + \tau_2 + \mu_2)C, \\ \dfrac{dR}{dt} = \phi_1 X + \phi_2 B + \phi_3 C - (\psi_1 + \psi_2 + \tau_3)R, \\ \dfrac{dE}{dt} = \tau_1 B + \tau_2 C + \tau_3 R - \mu_3 E. \end{cases} \tag{2.1}$$

The non-negative initial conditions are:

$$\begin{aligned} &X(0) = X_0 \geq 0, \qquad B(0) = B_0 \geq 0, \qquad C(0) = C_0 \geq 0, \\ &R(0) = R_0 \geq 0, \qquad E(0) = E_0 \geq 0. \end{aligned} \tag{2.2}$$

In System (2.1), Λ_1, Λ_2, and Λ_3 denote recruitment into the early-stage, Stage III and Stage IV compartments, respectively. Patients leave the early-stage class X through recovery at a rate $\phi_1 X$ or

progression to Stage III at rate $\beta_2 X$. The Stage III class B receives newly recruited patients, progressing patients from X, and relapsed patients from the recovered class R at a rate $\psi_1 R$. Its outflows are recovery, progression to stage IV, cardiotoxicity and cancer-related mortality, represented by $\phi_2 B$, $\beta_1 B$, $\tau_1 B$, and $\mu_1 B$, respectively. Similarly, Stage IV patients in C are recruited directly, arrive from B, or relapse from R at a rate $\psi_2 R$. Patients leave C through recovery, cardiotoxicity or cancer-related mortality at the rates $\phi_3 C$, $\tau_2 C$, and $\mu_2 C$. The recovered compartment R receives patients from X, B and C, but patients may relapse to Stage III or IV or develop delayed cardiotoxicity. Finally, the cardiotoxicity compartment E receives patients from B, C, and R at the rates $\tau_1 B$, $\tau_2 C$, and $\tau_3 R$, while $\mu_3 E$ represents its mortality outflow. These parameters describe assumed transitions in the mathematical model and should not be interpreted as patient-specific clinical estimates.

The biological meanings of the parameters in System (2.1) are presented in Table 1. All parameters are assumed to be non-negative.

Table 1. Biological interpretation of the model parameters.

Symbol	Description
Λ_1	Recruitment rate of patients diagnosed with Stage I or Stage II breast cancer
Λ_2	Recruitment rate of patients diagnosed with Stage III breast cancer
Λ_3	Recruitment rate of patients diagnosed with stage-IV breast cancer
ϕ_1	Recovery rate of patients in the early-stage compartment X
ϕ_2	Recovery rate of patients in the Stage III compartment B
ϕ_3	Recovery rate of patients in the Stage IV compartment C
β_1	Disease progression rate from Stage III to stage IV
β_2	Disease progression rate from Stage I or II to Stage III
τ_1	Rate at which patients in B develop chemotherapy-associated cardiotoxicity
τ_2	Rate at which patients in C develop chemotherapy-associated cardiotoxicity
τ_3	Rate at which recovered patients develop delayed cardiotoxicity
μ_1	Cancer-related mortality rate among patients in B
μ_2	Cancer-related mortality rate among patients in C
μ_3	Mortality rate among patients in the cardiotoxicity class E
ψ_1	Relapse rate from R to Stage III
ψ_2	Relapse rate from R to Stage IV

2.1. Numerical parameter settings and scope of calibration

The executable model parameter configuration supplied for the experiments specifies the following numerical coefficients. The above values are stored in the `Parameters.py` file and correspond to the coefficients of Eq (2.1). These values are reported to make the simulations reproducible; the available

files do not establish estimation at the level of individual patients, Saudi clinical calibration, measurement units, or uncertainty intervals. Thus the numbers should be interpreted as illustrative computational settings, not validated clinical rates.

Table 2. Numerical values of parameters recorded in the model's coding [39].

Parameter	Value	Parameter	Value
Λ_1	14000	Λ_2	80
Λ_3	90	β_1	0.0100
β_2	0.0340	ϕ_1	0.0300
ϕ_2	0.4000	ϕ_3	0.0100
ψ_1	0.0300	ψ_2	0.3000
τ_1	0.0300	τ_2	0.1000
τ_3	0.2000	μ_1	0.0256
μ_2	0.0256	μ_3	0.0256

The terminal time is $T = 100$ with $N = 1024$ steps ($\Delta t = 100/1024 = 0.09765625$). For the stochastic system solution $\sigma_i = 1$, where the noise is activated in the system and by $\sigma_i = 0$, the system becomes deterministic.

2.2. Stochastic extension

In real-world scenario, deterministic models cannot represent the complicated dynamics caused by environmental and demographic variation occurring in real life. On the other hand, the stochastic modeling gives an accurate description of breast cancer's dynamics in terms of biological variation, randomness, and uncertainty in the process of tumor development and therapy resistance. The current section presents the extension of the deterministic breast cancer model (2.1), which is known as stochastic model.

In System (2.3), $W_i(t)$, for $i = 1, \ldots, 5$, are independent standard Wiener processes, also known as Brownian motions. Their increments satisfy

$$dW_i(t) \sim \mathcal{N}(0, dt), \qquad i = 1, \ldots, 5,$$

meaning that each increment is normally distributed with a mean of zero and a variance dt. The stochastic intensity are represent by the parameters σ_i, namely σ_1, σ_2, σ_3, σ_4, and σ_5.

$$\begin{cases} dX = [\Lambda_1 - (\phi_1 + \beta_2)X]\,dt + \sigma_1 X\,dW_1, \\ dB = [\Lambda_2 + \beta_2 X + \psi_1 R - (\phi_2 + \beta_1 + \tau_1 + \mu_1)B]\,dt + \sigma_2 B\,dW_2, \\ dC = [\Lambda_3 + \beta_1 B + \psi_2 R - (\phi_3 + \tau_2 + \mu_2)C]\,dt + \sigma_3 C\,dW_3, \\ dR = [\phi_1 X + \phi_2 B + \phi_3 C - (\psi_1 + \psi_2 + \tau_3)R]\,dt + \sigma_4 R\,dW_4, \\ dE = [\tau_3 R + \tau_2 C + \tau_1 B - \mu_3 E]\,dt + \sigma_5 E\,dW_5. \end{cases} \tag{2.3}$$

3. Qualitative analysis of the deterministic and stochastic models

In this section, $\mathbf{Y}$ shows the five-dimensional state vector, the constant coefficient matrix of the deterministic drift is represented by A, and the recruitment vector is represented by $\mathbf{b}$. The diffusion

term works component-wise in the stochastic model. The following results hold for mathematical solutions under the assumed parameter conditions.

The current section addresses basic considerations on well-posedness, equilibrium, and expectation dynamics for the stochastic model. Existence proofs follow the standard techniques is briefly presented [40].

For notational convenience, define

$$\begin{aligned}\Omega_1 &= \phi_1 + \beta_2,\\ \Omega_2 &= \phi_2 + \beta_1 + \tau_1 + \mu_1,\\ \Omega_3 &= \phi_3 + \tau_2 + \mu_2,\\ \Omega_4 &= \psi_1 + \psi_2 + \tau_3.\end{aligned} \tag{3.1}$$

Let

$$A = \begin{pmatrix} -\Omega_1 & 0 & 0 & 0 & 0\\ \beta_2 & -\Omega_2 & 0 & \psi_1 & 0\\ 0 & \beta_1 & -\Omega_3 & \psi_2 & 0\\ \phi_1 & \phi_2 & \phi_3 & -\Omega_4 & 0\\ 0 & \tau_1 & \tau_2 & \tau_3 & -\mu_3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} \Lambda_1\\ \Lambda_2\\ \Lambda_3\\ 0\\ 0 \end{pmatrix}. \tag{3.2}$$

3.1. Existence and uniqueness of the deterministic solution

We investigate the existence and uniqueness of the model (2.1). To perform this, we present the model (2.1) in the form given by

$$\begin{aligned} X'(t) &= f_1(t, X, B, C, R, E),\\ B'(t) &= f_2(t, X, B, C, R, E),\\ C'(t) &= f_3(t, X, B, C, R, E),\\ R'(t) &= f_4(t, X, B, C, R, E),\\ E'(t) &= f_5(t, X, B, C, R, E).\end{aligned} \tag{3.3}$$

We define the norm given by

$$\|X\|_\infty = \sup_{t\in U} |X(t)|, \tag{3.4}$$

where $U = [0, T]$. Here, it is assumed that $X(t)$, $B(t)$, $C(t)$, $R(t)$, and $E(t)$ are bounded in $[0, T]$ for any $t \in [0, T]$, and $S_1, \ldots, S_5$ exist such that $\|X\|_\infty < S_1$, $\|B\|_\infty < S_2$, $\|C\|_\infty < S_3$, $\|R\|_\infty < S_4$, and $\|E\|_\infty < S_5$. We have to prove that $f_1, \ldots, f_5$ are bounded.

$$\begin{aligned} |f_1(t, X, B, C, R, E)| &= |\Lambda_1 - (\phi_1 + \beta_2)X|\\ &\le \Lambda_1 + (\phi_1 + \beta_2)|X|\\ &\le \Lambda_1 + (\phi_1 + \beta_2)\sup_t |X|\\ &\le \Lambda_1 + (\phi_1 + \beta_2)\|X\|_\infty\\ &\le \Lambda_1 + (\phi_1 + \beta_2)S_1 < \infty,\end{aligned} \tag{3.5}$$

where $t \in [0, T]$. We follow the method above and can easily show the rest as follows:

$$\begin{aligned} |f_2(t, X, B, C, R, E)| &= |\Lambda_2 + \beta_2 X + \psi_1 R - \Phi_2 B|\\ &\le \Lambda_2 + \beta_2 S_1 + \psi_1 S_4 + \Phi_2 S_2 < \infty.\end{aligned} \tag{3.6}$$

$$|f_3(t,X,B,C,R,E)| = |\Lambda_3 + \beta_1 B + \psi_2 R - \Phi_3 C| \leq \Lambda_3 + \beta_1 S_2 + \psi_2 S_4 + \Phi_3 S_3 < \infty. \tag{3.7}$$

$$|f_4(t,X,B,C,R,E)| = |\phi_1 X + \phi_2 B + \phi_3 C - (\psi_1 + \psi_2 + \tau_3)R| \leq \phi_1 S_1 + \phi_2 S_2 + \phi_3 S_3 + \Phi_4 S_4 < \infty. \tag{3.8}$$

$$|f_5(t,X,B,C,R,E)| = |\tau_3 R + \tau_2 C + \eta B - \mu_3 E| \leq \tau_3 S_4 + \tau_2 S_3 + \eta S_2 + \mu_3 S_5 < \infty, \tag{3.9}$$

where $\Phi_2 = (\phi_2 + \beta_1 + \tau_1 + \mu_1)$, $\Phi_3 = (\phi_3 + \tau_2 + \mu_2)$, and $\Phi_4 = (\psi_1 + \psi_2 + \tau_3)$.

Thus, $X(t)$, $B(t)$, $C(t)$, $R(t)$, and $E(t)$ are bounded, and then $l_1, \ldots, l_5$ exist such that

$$\begin{aligned} &\sup_t |f_1(t,X,B,C,R,E)| < l_1, \\ &\sup_t |f_2(t,X,B,C,R,E)| < l_2, \\ &\sup_t |f_3(t,X,B,C,R,E)| < l_3, \\ &\sup_t |f_4(t,X,B,C,R,E)| < l_4, \\ &\sup_t |f_5(t,X,B,C,R,E)| < l_5. \end{aligned} \tag{3.10}$$

Further, we have to show

$$\begin{aligned} &|f_1(t,X_1,B,C,R,E) - f_1(t,X_2,B,C,R,E)| \\ &= |\Lambda_1 - (\phi_1 + \beta_2)X_1 - \Lambda_1 + (\phi_1 + \beta_2)X_2| \\ &< (\phi_1 + \beta_2)|X_1 - X_2| \\ &< \overline{\varpi}_1 |X_1 - X_2|, \end{aligned} \tag{3.11}$$

where $\overline{\varpi}_1 = (\phi_1 + \beta_2)$. Following the same approach, we show

$$\begin{aligned} &|f_2(t,X,B_1,C,R,E) - f_2(t,X,B_2,C,R,E)| < \overline{\varpi}_2 |B_1 - B_2|, \\ &|f_3(t,X,B,C_1,R,E) - f_3(t,X,B,C_2,R,E)| < \overline{\varpi}_3 |C_1 - C_2|, \\ &|f_4(t,X,B,C,R_1,E) - f_4(t,X,B,C,R_2,E)| < \overline{\varpi}_4 |R_1 - R_2|, \\ &|f_5(t,X,B,C,R,E_1) - f_5(t,X,B,C,R,E_2)| < \overline{\varpi}_5 |E_1 - E_2|, \end{aligned} \tag{3.12}$$

where $\overline{\varpi}_2 = (\phi_2 + \beta_1 + \tau_1 + \mu_1)$, $\overline{\varpi}_3 = (\phi_3 + \tau_2 + \mu_2)$, $\overline{\varpi}_4 = (\psi_1 + \psi_2 + \tau_3)$, and $\overline{\varpi}_5 = \mu_3$.

This uniqueness and existence of the solution of Model (2.1) indicate that the model has a unique system of solutions.

3.2. Equilibrium of the deterministic model

Let

$$\mathcal{P}^* = (X^*, B^*, C^*, R^*, E^*) \tag{3.13}$$

be an equilibrium of model (2.1). Setting all time derivatives equal to zero gives

$$X^* = \frac{\Lambda_1}{\Omega_1}. \tag{3.14}$$

Define

$$H_1 = \Lambda_2 + \beta_2 X^*, \qquad H_2 = \phi_1 X^*. \tag{3.15}$$

The variables B^*, C^*, and R^* satisfy

$$\begin{pmatrix} \Omega_2 & 0 & -\psi_1 \\ -\beta_1 & \Omega_3 & -\psi_2 \\ -\phi_2 & -\phi_3 & \Omega_4 \end{pmatrix} \begin{pmatrix} B^* \\ C^* \\ R^* \end{pmatrix} = \begin{pmatrix} H_1 \\ \Lambda_3 \\ H_2 \end{pmatrix}. \tag{3.16}$$

The determinant of the coefficient matrix is

$$\Delta = \Omega_2\Omega_3\Omega_4 - \Omega_2\psi_2\phi_3 - \psi_1\beta_1\phi_3 - \psi_1\Omega_3\phi_2. \tag{3.17}$$

If $\Delta \neq 0$, the equilibrium is unique and

$$R^* = \frac{\Omega_2\Omega_3 H_2 + (\Omega_3\phi_2 + \beta_1\phi_3)H_1 + \Omega_2\phi_3\Lambda_3}{\Delta}, \tag{3.18}$$

$$B^* = \frac{H_1 + \psi_1 R^*}{\Omega_2}, \qquad C^* = \frac{\Lambda_3 + \beta_1 B^* + \psi_2 R^*}{\Omega_3}, \tag{3.19}$$

and

$$E^* = \frac{\tau_1 B^* + \tau_2 C^* + \tau_3 R^*}{\mu_3}, \tag{3.20}$$

when all parameters are non-negative, $\mu_3 > 0$, and $\Delta > 0$, all components of $\mathcal{P}^*$ are non-negative.

3.3. Existence, uniqueness, and positivity of the stochastic solution

For

$$\mathbf{W} = (W_1, \ldots, W_5)^{\mathsf{T}}$$

consisting of independent standard Wiener processes, the stochastic model is written as

$$d\mathbf{Y}(t) = (A\mathbf{Y}(t) + \mathbf{b})dt + G(\mathbf{Y}(t))\, d\mathbf{W}(t), \tag{3.21}$$

where

$$G(\mathbf{Y}) = \operatorname{diag}(\sigma_1 X, \sigma_2 B, \sigma_3 C, \sigma_4 R, \sigma_5 E). \tag{3.22}$$

Both $A\mathbf{Y} + \mathbf{b}$ and $G(\mathbf{Y})$ are globally Lipschitz and satisfy linear-growth bounds. Therefore, for a square-integrable initial state, the standard Itô existence and uniqueness theorem [40] provides a unique global strong solution.

Furthermore, when all coefficients and the initial state are non-negative, the drift of each compartment at its zero boundary is non-negative, and its own diffusion coefficient is zero. For example,

$$f_X|_{X=0} = \Lambda_1 \geq 0$$

and

$$f_E|_{E=0} = \tau_1 B + \tau_2 C + \tau_3 R \geq 0;$$

the remaining boundaries satisfy the same property by Eq (2.1). These inward-pointing drift and boundary-degenerate diffusion conditions preserve the non-negative orthant almost surely.

3.4. Expected value dynamics of the stochastic model

As the stochastic disturbances can be written as Ito integrals with zero expectation, then it follows that the mean state $\mathbf{m}(t) = \mathbb{E}[\mathbf{Y}(t)]$ satisfies an affine drift deterministic system as follows:

$$\mathbf{m}'(t) = A\mathbf{m}(t) + \mathbf{b}, \qquad \mathbf{m}(t) = e^{At}\mathbf{m}(0) + \int_0^t e^{A(t-s)}\mathbf{b}\, ds. \tag{3.23}$$

If $\Omega_1 > 0$, $\mu_3 > 0$, and the Routh–Hurwitz conditions hold, then

$$\lim_{t\to\infty} \mathbb{E}[\mathbf{Y}(t)] = -A^{-1}\mathbf{b} = \mathcal{P}^*. \tag{3.24}$$

This follows directly from $e^{At} \to 0$ and is a statement about the mean, not about the convergence of individual stochastic paths.

Finally, define the total population

$$N(t) = X(t) + B(t) + C(t) + R(t) + E(t).$$

Summing the five stochastic equations yields

$$\begin{aligned} dN(t) = {} & [\Lambda_1 + \Lambda_2 + \Lambda_3 - \mu_1 B(t) - \mu_2 C(t) - \mu_3 E(t)]\, dt \\ & + \sigma_1 X(t)\, dW_1(t) + \sigma_2 B(t)\, dW_2(t) + \sigma_3 C(t)\, dW_3(t) \\ & + \sigma_4 R(t)\, dW_4(t) + \sigma_5 E(t)\, dW_5(t). \end{aligned} \tag{3.25}$$

This balance shows that all transfers between compartments cancel from the total population equation.

4. Iterative numerical solution via the Legendre spectral collocation method

This section present the application of the LSCM to numerically solve the stochastic model (2.3) and generate reference data for HQNN [41–43]. The LSCM has been extensively used in order to numerically solve the differential and integral equations. For any function $h(\xi)$, defined in the standard domain $[-1, 1]$, the truncated Legendre expansion up to degree M is written by

$$h_M(\xi) = \sum_{\iota=0}^{M} h_\iota K_\iota(\xi), \tag{4.1}$$

where $K_\iota(\xi)$ denotes the Legendre polynomial of degree ι, while h_ι represents the corresponding unknown expansion coefficient. The collocation points satisfy

$$-1 = \xi_0 < \xi_1 < \cdots < \xi_M = 1.$$

An explicit representation of the Legendre polynomial is

$$K_\iota(\xi) = \frac{1}{2^\iota} \sum_{j=0}^{\lfloor \iota/2 \rfloor} (-1)^j \frac{(2\iota - 2j)!}{j!(\iota - j)!(\iota - 2j)!} \xi^{\iota - 2j}, \qquad \iota = 0, \ldots, M. \tag{4.2}$$

Here, the greatest-integer function is defined by

$$\left\lfloor \frac{\iota}{2} \right\rfloor = \begin{cases} \dfrac{\iota}{2}, & \text{if } \iota \text{ is even,} \\ \dfrac{\iota - 1}{2}, & \text{if } \iota \text{ is odd.} \end{cases} \tag{4.3}$$

The stochastic system used below is the model in Eq (2.3). Integrating System (2.3) over the interval $[0, t]$ gives

$$\begin{cases} X(t) = X_0 + \displaystyle\int_0^t \left[\Lambda_1 - (\phi_1 + \beta_2)X(\zeta)\right] d\zeta + \int_0^t \sigma_1 X(\zeta)\, dW_1(\zeta), \\ B(t) = B_0 + \displaystyle\int_0^t \left[\Lambda_2 + \beta_2 X(\zeta) + \psi_1 R(\zeta) - (\phi_2 + \beta_1 + \tau_1 + \mu_1)B(\zeta)\right] d\zeta \\ \qquad + \displaystyle\int_0^t \sigma_2 B(\zeta)\, dW_2(\zeta), \\ C(t) = C_0 + \displaystyle\int_0^t \left[\Lambda_3 + \beta_1 B(\zeta) + \psi_2 R(\zeta) - (\phi_3 + \tau_2 + \mu_2)C(\zeta)\right] d\zeta \\ \qquad + \displaystyle\int_0^t \sigma_3 C(\zeta)\, dW_3(\zeta), \\ R(t) = R_0 + \displaystyle\int_0^t \left[\phi_1 X(\zeta) + \phi_2 B(\zeta) + \phi_3 C(\zeta) - (\psi_1 + \psi_2 + \tau_3)R(\zeta)\right] d\zeta \\ \qquad + \displaystyle\int_0^t \sigma_4 R(\zeta)\, dW_4(\zeta), \\ E(t) = E_0 + \displaystyle\int_0^t \left[\tau_1 B(\zeta) + \tau_2 C(\zeta) + \tau_3 R(\zeta) - \mu_3 E(\zeta)\right] d\zeta \\ \qquad + \displaystyle\int_0^t \sigma_5 E(\zeta)\, dW_5(\zeta). \end{cases} \tag{4.4}$$

To transform the integration interval $[0, t]$ into the standard interval $[-1, 1]$, we introduce the linear transformation

$$\zeta = \frac{t}{2}(v + 1), \qquad -1 \le v \le 1, \tag{4.5}$$

for which

$$d\zeta = \frac{t}{2}\, dv.$$

Applying this transformation to Eq (4.4) gives

$$
\begin{aligned}
X(t) &= X_0 + \frac{t}{2}\int_{-1}^{1}\left[\Lambda_1 - (\phi_1+\beta_2)X\left(\frac{t}{2}(v+1)\right)\right]dv \\
&\quad + \int_{-1}^{1}\sigma_1 X\left(\frac{t}{2}(v+1)\right)dW_1\left(\frac{t}{2}(v+1)\right), \\
B(t) &= B_0 + \frac{t}{2}\int_{-1}^{1}\left[\Lambda_2 + \beta_2 X\left(\frac{t}{2}(v+1)\right) + \psi_1 R\left(\frac{t}{2}(v+1)\right)\right. \\
&\quad \left. - (\phi_2+\beta_1+\tau_1+\mu_1)B\left(\frac{t}{2}(v+1)\right)\right]dv \\
&\quad + \int_{-1}^{1}\sigma_2 B\left(\frac{t}{2}(v+1)\right)dW_2\left(\frac{t}{2}(v+1)\right), \\
C(t) &= C_0 + \frac{t}{2}\int_{-1}^{1}\left[\Lambda_3 + \beta_1 B\left(\frac{t}{2}(v+1)\right) + \psi_2 R\left(\frac{t}{2}(v+1)\right)\right. \\
&\quad \left. - (\phi_3+\tau_2+\mu_2)C\left(\frac{t}{2}(v+1)\right)\right]dv \\
&\quad + \int_{-1}^{1}\sigma_3 C\left(\frac{t}{2}(v+1)\right)dW_3\left(\frac{t}{2}(v+1)\right), \\
R(t) &= R_0 + \frac{t}{2}\int_{-1}^{1}\left[\phi_1 X\left(\frac{t}{2}(v+1)\right) + \phi_2 B\left(\frac{t}{2}(v+1)\right)\right. \\
&\quad \left. + \phi_3 C\left(\frac{t}{2}(v+1)\right) - (\psi_1+\psi_2+\tau_3)R\left(\frac{t}{2}(v+1)\right)\right]dv \\
&\quad + \int_{-1}^{1}\sigma_4 R\left(\frac{t}{2}(v+1)\right)dW_4\left(\frac{t}{2}(v+1)\right), \\
E(t) &= E_0 + \frac{t}{2}\int_{-1}^{1}\left[\tau_1 B\left(\frac{t}{2}(v+1)\right) + \tau_2 C\left(\frac{t}{2}(v+1)\right)\right. \\
&\quad \left. + \tau_3 R\left(\frac{t}{2}(v+1)\right) - \mu_3 E\left(\frac{t}{2}(v+1)\right)\right]dv \\
&\quad + \int_{-1}^{1}\sigma_5 E\left(\frac{t}{2}(v+1)\right)dW_5\left(\frac{t}{2}(v+1)\right).
\end{aligned} \tag{4.6}
$$

Let

$$-1 = v_0 < v_1 < \cdots < v_M = 1$$

be the Legendre-Gauss-Lobatto points, and define the corresponding mapped time points by

$$\zeta_j(t) = \frac{t}{2}(v_j+1), \qquad j = 0, \ldots, M. \tag{4.7}$$

For compactness, introduce

$$
\begin{aligned}
&X_j = X(\zeta_j(t)), \quad B_j = B(\zeta_j(t)), \quad C_j = C(\zeta_j(t)), \\
&R_j = R(\zeta_j(t)), \quad E_j = E(\zeta_j(t)).
\end{aligned}
$$

The semi discrete spectral equations are then expressed as

$$X(t) = X_0 + \frac{t}{2}\sum_{j=0}^{M}\Big[\Lambda_1 - (\phi_1 + \beta_2)X_j\Big]\omega_j + \sum_{j=0}^{M-1}\sigma_1 X_j \Delta W_{1,j},$$

$$B(t) = B_0 + \frac{t}{2}\sum_{j=0}^{M}\Big[\Lambda_2 + \beta_2 X_j + \psi_1 R_j - (\phi_2 + \beta_1 + \tau_1 + \mu_1)B_j\Big]\omega_j + \sum_{j=0}^{M-1}\sigma_2 B_j \Delta W_{2,j},$$

$$C(t) = C_0 + \frac{t}{2}\sum_{j=0}^{M}\Big[\Lambda_3 + \beta_1 B_j + \psi_2 R_j - (\phi_3 + \tau_2 + \mu_2)C_j\Big]\omega_j + \sum_{j=0}^{M-1}\sigma_3 C_j \Delta W_{3,j},$$

$$R(t) = R_0 + \frac{t}{2}\sum_{j=0}^{M}\Big[\phi_1 X_j + \phi_2 B_j + \phi_3 C_j - (\psi_1 + \psi_2 + \tau_3)R_j\Big]\omega_j + \sum_{j=0}^{M-1}\sigma_4 R_j \Delta W_{4,j},$$

$$E(t) = E_0 + \frac{t}{2}\sum_{j=0}^{M}\Big[\tau_1 B_j + \tau_2 C_j + \tau_3 R_j - \mu_3 E_j\Big]\omega_j + \sum_{j=0}^{M-1}\sigma_5 E_j \Delta W_{5,j}. \tag{4.8}$$

The Legendre-Gauss-Lobatto quadrature weights corresponding to the collocation points are given by

$$\omega_j = \frac{2}{M(M+1)\Big[K_M(v_j)\Big]^2}, \qquad j = 0, \ldots, M. \tag{4.9}$$

The Brownian increments appearing in Eq (4.8) are defined by

$$\Delta W_{i,j} = W_i(\zeta_{j+1}(t)) - W_i(\zeta_j(t)), \qquad i = 1, \ldots, 5, \tag{4.10}$$

and satisfy

$$\Delta W_{i,j} \sim \mathcal{N}\left(0,\, \Delta\zeta_j\right), \qquad \Delta\zeta_j = \zeta_{j+1}(t) - \zeta_j(t). \tag{4.11}$$

For numerical implementation, the increments can be generated from

$$\Delta W_{i,j} = \sqrt{\Delta\zeta_j}\,\mathcal{Z}_{i,j}, \qquad \mathcal{Z}_{i,j} \sim \mathcal{N}(0,1), \tag{4.12}$$

where all random variables $\mathcal{Z}_{i,j}$ are assumed to be mutually independent.

Using the Legendre polynomial representation in Eq (4.1), the approximations of the five state variables are written as

$$\begin{aligned}
X_M(t) &= \sum_{\vartheta=0}^{M} X_\vartheta K_\vartheta\left(\frac{2t}{T} - 1\right), \quad B_M(t) = \sum_{\vartheta=0}^{M} B_\vartheta K_\vartheta\left(\frac{2t}{T} - 1\right),\\
C_M(t) &= \sum_{\vartheta=0}^{M} C_\vartheta K_\vartheta\left(\frac{2t}{T} - 1\right), \quad R_M(t) = \sum_{\vartheta=0}^{M} R_\vartheta K_\vartheta\left(\frac{2t}{T} - 1\right),\\
E_M(t) &= \sum_{\vartheta=0}^{M} E_\vartheta K_\vartheta\left(\frac{2t}{T} - 1\right).
\end{aligned} \tag{4.13}$$

Substituting the approximations in Eq (4.13) into the semi-discrete system (4.8) and imposing the resulting relations at the collocation points generate the algebraic system for the unknown Legendre coefficients. For each prescribed realization of the Brownian increments, the unknown coefficients are

$$\{X_\vartheta, B_\vartheta, C_\vartheta, R_\vartheta, E_\vartheta\}_{\vartheta=0}^{M}.$$

Therefore, the approximation contains $5M + 5$ unknown coefficients.

The initial conditions are imposed at $t = 0$. Because

$$\frac{2(0)}{T} - 1 = -1,$$

they take the form

$$\begin{aligned} &\sum_{\vartheta=0}^{M} X_\vartheta K_\vartheta(-1) = X_0, \quad \sum_{\vartheta=0}^{M} B_\vartheta K_\vartheta(-1) = B_0, \\ &\sum_{\vartheta=0}^{M} C_\vartheta K_\vartheta(-1) = C_0, \quad \sum_{\vartheta=0}^{M} R_\vartheta K_\vartheta(-1) = R_0, \\ &\sum_{\vartheta=0}^{M} E_\vartheta K_\vartheta(-1) = E_0. \end{aligned} \tag{4.14}$$

The five initial conditions in Eq (4.14), together with the $5M$ collocation equations obtained from Eq (4.8), form a square algebraic system of $5M + 5$ equations for the $5M + 5$ unknown coefficients

$$X_\vartheta, B_\vartheta, C_\vartheta, R_\vartheta, E_\vartheta, \qquad \vartheta = 0, \ldots, M.$$

5. The HQNN architecture

For reproducibility, we distinguish the formal LSCM in Section 4 from the supplied executable reference-generation code. The `numericalsolver.py` is used to generate the deterministic and stochastic data; the HQNN ablation program loads these trajectories from CSV files. The saved reference grids contain 1025 time points over $[0, 100]$. The original reconstruction trains on 129 approximately equally spaced time samples and displays predictions over the complete grid; this is trajectory interpolation and includes the time points used during training. The later ablation instead reserves an interior block for testing ($40 - 60\%$ of the trajectory), a separate block for validation ($20 - 30\%$), and estimates scaling coefficients from training values only.

5.1. Choice of numerical and hybrid learning methods

The numerical and learning procedures have different purposes in this investigation. The numerical scheme generates reference trajectories for the mathematical model, whereas the HQNN learns a time-dependent approximation from the selected numerical values. It is therefore important to distinguish the accuracy of the reference solution from the reconstruction accuracy of the trained surrogate.

For smooth deterministic problems, LSCM may provide accurate global polynomial representations with comparatively few expansion coefficients when the solution is sufficiently regular. This is the motivation for describing its formulation in Section 4. Since stochastic sample paths are not generally

smooth, the convergence behavior expected of polynomial approximations to smooth deterministic functions cannot be assumed for the stochastic model without a dedicated test.

The HQNN was investigated as a supervised multi output surrogate because its time encoding, trainable quantum rotations, and Pauli expectations provide a nonlinear feature representation shared by the five biological compartments. After fitting, the same trained network can be evaluated at the requested time coordinates. Nevertheless, the training stage requires repeated quantum circuit simulation and does not replace the numerical generation of training targets. In the present work, the circuit is evaluated with a classical state-vector simulator, for which the state dimension grows at 2^n with the number of qubits n. No measured training time benefit, quantum speedup, or higher predictive accuracy relative to conventional neural networks is established by the available experiments. A comparative analysis of the reference numerical and neural network methods used in the present study is presented in Table 3.

Table 3. Comparison of the numerical and neural network methods used in the present study.

Method	Purpose and advantages	Limitations
Deterministic LSCM	The method is used to obtain numerical solutions of the deterministic model through Legendre polynomial approximations. It can provide accurate results for sufficiently smooth solutions.	The numerical accuracy depends on the selected polynomial order. A separate convergence study has not been carried out.
Stochastic LSCM	The method is used to solve the stochastic model by including Brownian increments in the numerical formulation. The resulting solutions are used as reference data for the HQNN.	The accuracy of the stochastic solutions requires further investigation through comparison with standard numerical methods and convergence tests.
Hybrid quantum neural network	The HQNN is used to approximate the five compartmental solutions simultaneously. It combines quantum features with a classical neural network to learn the numerical trajectories.	The method requires reference data for training. Its performance depends on the training process, and no quantum computational advantage has been established.

The model parameters, initial conditions, and the temporal discretization are first specified for either the deterministic or stochastic experiment. The resulting reference trajectory is scaled componentwise, and 129 approximately equidistant observations are selected for training. Normalized time is processed by the batched five-qubit parameterized quantum circuit, which generates 15 Pauli expectation features for the classical 15-24-24-5 regression network. Quantum and

classical parameters are optimized jointly by minimizing a composite loss comprising the trajectory-fitting error and an initial condition penalty. After training, the model is evaluated on the complete 1025-point time grid, transformed back to the physical scale, and assessed using trajectory plots and componentwise error metrics. The computation process used in the numerical simulation can be summarized by Figure 1. For the deterministic system, the reference dataset is obtained according to the dynamic equation. For the stochastic simulation, a reference trajectory is obtained by solving the stochastic system. Once the reference trajectory is acquired, all computations follow the same scaling, subsampling, quantum features acquisition, classical regression modeling, and optimization processes.

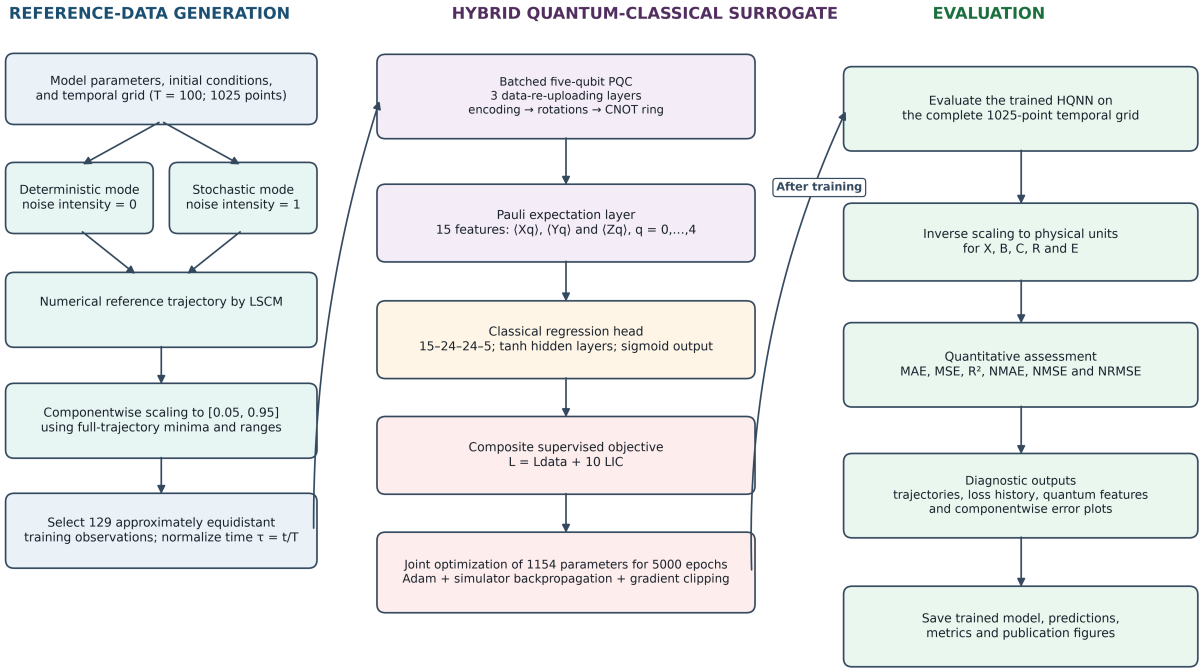

Figure 1. Workflow of the supervised hybrid quantum classical surrogate.

The process of training is carried out by running the quantum circuit for all 129 normalized temporal inputs during each epoch in a batched manner. The resulting 15 Pauli expectations at each time step are fed into the classical regression network. Gradients of the quantum rotation angles are calculated by back propagating through the state vector simulation. The gradients of the classical weights and biases are calculated by automatic differentiation. Hence, the optimization of the quantum rotation angles and classical regression network is performed simultaneously using the Adam optimizer. After a maximum of 5000 training epochs (with early stopping in the later ablation experiments), the results are predicted at each time step on the entire temporal domain, followed by normalization. The end-to-end HQNN algorithm for the cancer cardio-toxicity model is presented in Algorithm 1.

Algorithm 1 End-to-end HQNN algorithm for the cancer cardio-toxicity model

Input: Model parameters $\boldsymbol{p}$, initial state $\boldsymbol{y}_0$, time interval $[0, T]$, time steps N, training points N_{tr}, epochs K, learning rate η_0, initial condition (IC) weight λ_{IC}, noise indicator ν, diffusion vector $\boldsymbol{\sigma}$, and seed s
Output: Trained parameters $\boldsymbol{\Theta}^*$, predicted trajectory $\widehat{\boldsymbol{Y}}$, and error metrics
Step 1: Generate the reference trajectory
Provide a numerical reference dataset
Step 2: Normalize the data
for $j = 1, \ldots, 5$ **do**
    Compute $m_j = \min_n y_{n,j}^{\mathrm{ref}}$, $M_j = \max_n y_{n,j}^{\mathrm{ref}}$, and $d_j = \max(M_j - m_j, 10^{-12})$
    $\widetilde{y}_{n,j} = 0.05 + 0.90(y_{n,j}^{\mathrm{ref}} - m_j)/d_j$
end for
Select N_{tr} approximately equidistant samples and set $\tau_i = t_i/T$
Step 3: Construct the HQNN
Initialize five qubits, three data reuploading layers, and $\boldsymbol{\theta} \in \mathbb{R}^{3\times5\times3}$
In each quantum layer, apply $R_Y(\omega_q\tau_i)$, $R_Z(\omega_q\tau_i/2)$, trainable $R_X R_Y R_Z$ rotations, and the circular CNOT ring
Measure Pauli-X, Pauli-Y, and Pauli-Z observable to obtain $z_i^{(Q)} \in [-1, 1]^{15}$
Construct the classical regression head 15–24–24–5 with tanh, and sigmoid activations
Initialize $\boldsymbol{\Theta} = (\boldsymbol{\theta}, \boldsymbol{\phi})$, Adam, and the reduce-on-plateau scheduler
Step 4: Train the HQNN
for $k = 1, \ldots, K$ **do**
    Evaluate all τ_i in one batched HQNN forward pass
    $\widehat{\widetilde{\boldsymbol{y}}}_i = C_{\boldsymbol{\phi}}(Q_{\boldsymbol{\theta}}(\tau_i))$
    Compute $\mathcal{L}_{\mathrm{data}} = (5N_{\mathrm{tr}})^{-1} \sum_i \|\widehat{\widetilde{\boldsymbol{y}}}_i - \widetilde{\boldsymbol{y}}_i\|_2^2$
    Compute $\mathcal{L}_{\mathrm{IC}} = 5^{-1}\|\widehat{\widetilde{\boldsymbol{y}}}(0) - \widetilde{\boldsymbol{y}}_0\|_2^2$
    Set $\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \lambda_{\mathrm{IC}}\mathcal{L}_{\mathrm{IC}}$
    Backpropagate $\mathcal{L}$ through the quantum and classical components
    Clip the gradient norm at 5 and update $\boldsymbol{\Theta}$ using Adam
    Update the learning-rate scheduler using $\mathcal{L}$
end for
Step 5: Full-grid prediction and post-processing
Set $\boldsymbol{\Theta}^* \leftarrow \boldsymbol{\Theta}$ and evaluate the trained HQNN at $\tau_n = t_n/T$, $n = 0, \ldots, N$
for $n = 0, \ldots, N$ **do**
    for $j = 1, \ldots, 5$ **do**
        $$\widehat{y}_{n,j} = m_j + \frac{d_j}{0.90}\left(\widehat{\widetilde{y}}_{n,j} - 0.05\right)$$
    end for
end for
Compute MAE, MSE, RMSE, Normalized MAE (NMAE), Normalized MSE (NMSE), Normalized RMSE (NRMSE), and R^2
Save $\boldsymbol{\Theta}^*$, $\widehat{\boldsymbol{Y}}$, metrics, and figures
Return: $\boldsymbol{\Theta}^*$, $\widehat{\boldsymbol{Y}}$, and the evaluation metrics

5.2. Problem formulation and end-to-end architecture

The HQNN is formulated as a supervised, multi output surrogate for the numerical trajectories generated by the five-compartment cancer cardiotoxicity model. Let

$$\boldsymbol{y}(t) = \begin{bmatrix} X(t) & B(t) & C(t) & R(t) & E(t) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^5 \tag{5.1}$$

denote the model's state at the physical time $t \in [0, T]$. The objective is to construct a parameterized approximation

$$\widehat{\boldsymbol{y}}(t; \boldsymbol{\Theta}) = \begin{bmatrix} \widehat{X}(t) & \widehat{B}(t) & \widehat{C}(t) & \widehat{R}(t) & \widehat{E}(t) \end{bmatrix}^{\mathrm{T}}, \tag{5.2}$$

where

$$\boldsymbol{\Theta} = (\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{5.3}$$

contains the trainable parameters of the parameterized quantum circuit (PQC), denoted $\boldsymbol{\theta}$, and those of the classical regression network, denoted $\boldsymbol{\phi}$.

The implemented HQNN is expressed as the composition

$$\widehat{\boldsymbol{y}}(t; \boldsymbol{\Theta}) = \mathcal{S}^{-1}\left[C_{\boldsymbol{\phi}}\left(Q_{\boldsymbol{\theta}}\left(\frac{t}{T}\right)\right)\right], \tag{5.4}$$

where

$$Q_{\boldsymbol{\theta}} : [0, 1] \longrightarrow [-1, 1]^{15} \tag{5.5}$$

is the parameterized quantum feature map,

$$C_{\boldsymbol{\phi}} : [-1, 1]^{15} \longrightarrow (0, 1)^5. \tag{5.6}$$

The equation above represents the classical multi output regression mapping, while $\mathcal{S}^{-1}$ is referred to as the inverse state-scaling transformation. In the current study, the mapping is conducted for the normalized time to 15 quantum expectation features, which are further scaled to produce five normalized states.

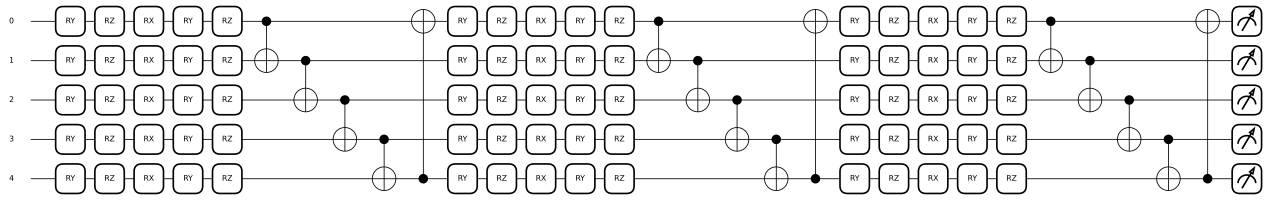

Figure 2. Architecture of the implemented five-qubit parameterized quantum circuit. Each of the three data reuploading layers comprises qubit-dependent temporal encoding through the R_Y and R_Z rotations; trainable R_X, R_Y; and R_Z rotations, and a circular CNOT entangling block. After the final layer, the expectation values of the Pauli-X, Pauli-Y, and Pauli-Z observables are evaluated on all five qubits, producing 15 real-valued quantum features. These features are supplied to a 15-24-24-5 classical regression network to obtain the five normalized state predictions.

For the PQC, we use the default qubit state vector simulator and set the interface to be PyTorch. We calculate the gradients of the loss concerning quantum parameters through the differentiation method $diff_method$ = "$backprop$". In this way, both the quantum and classical parameters are in one computational graph that allows updating both of them using the same optimization procedure. Temporal training batches are used in the evaluation process without requiring any for-loop of the time steps in Python.

We construct an objective based on the difference between HQNN prediction and target trajectories that were numerically calculated. In addition, there is no term associated with the residual of the governing differential equations within our objective function; therefore, the approach is named the supervised HQNN surrogate method rather than the physics-informed quantum neural network method.

5.3. Normalization of time and state variables

The physical time mapping on the unit interval is described by

$$\tau = \frac{t}{T}, \qquad 0 \le \tau \le 1, \tag{5.7}$$

where T denotes the terminal simulation time. This normalized scalar constitutes the input to the quantum circuit.

$$m_j = \min_{0\le n\le N} y_j(t_n), \tag{5.8}$$

$$M_j = \max_{0\le n\le N} y_j(t_n), \tag{5.9}$$

$$d_j = \max\left(M_j - m_j, 10^{-12}\right). \tag{5.10}$$

Here, $\{t_n\}_{n=0}^{N}$ is the reference time grid. The corresponding normalized target is

$$\widetilde{y}_j(t) = 0.05 + 0.90\frac{y_j(t) - m_j}{d_j}. \tag{5.11}$$

The lower bound imposed on d_j prevents division by zero for a constant or numerically invariant state.

The range [0.05,0.95] is chosen instead of [0,1] due to the application of the sigmoid activation function used by the classical output layer. Avoiding proximity to the asymptotes minimizes saturation by the sigmoid function. In order to transform the predicted value into its physical quantity, we utilize the transformation of the normalized output.

$$\widehat{y}_j(t) = m_j + \frac{d_j}{0.90}\left[\widehat{\widetilde{y}}_j(t/T) - 0.05\right]. \tag{5.12}$$

5.4. Parameterized quantum feature map

The quantum register contains $n = 5$ qubits and is initialized in the computational zero state,

$$|\psi^{(0)}\rangle = |0\rangle^{\otimes 5} = |00000\rangle. \tag{5.13}$$

Its Hilbert-space dimension is therefore $2^5 = 32$.

The circuit contains $L = 3$ data reuploading layers. A distinct, non-trainable frequency is assigned to each qubit as follows:

$$\boldsymbol{\omega} = \pi \begin{bmatrix} 0.70 & 1.35 & 2.00 & 2.65 & 3.30 \end{bmatrix}^{\mathrm{T}}. \tag{5.14}$$

For the qubit $q \in \{0, 1, 2, 3, 4\}$, the normalized temporal input is encoded using

$$R_Y(\omega_q \tau) \quad \text{and} \quad R_Z\left(\frac{\omega_q \tau}{2}\right). \tag{5.15}$$

The following results have been derived regarding the qubits' angular frequency-dependent trigonometric analysis. In qubits, there is more than one interpretation of identical scalar information based on their angular frequency. As the frequency decreases, the angular rotations decrease; otherwise, they increase with an increase in the time variation.

Each qubit experiences three trainable rotations after temporal encoding. For layer ℓ and qubit q, the local variational operator is

$$V_{\ell q} = R_Z\left(\theta^z_{\ell q}\right) R_Y\left(\theta^y_{\ell q}\right) R_X\left(\theta^x_{\ell q}\right). \tag{5.16}$$

The process described in (5.16) follows the operator rule in which the operation on the far-right side acts first. After each variational layer, the circular CNOT layout follows.

$$0 \to 1, \qquad 1 \to 2, \qquad 2 \to 3, \qquad 3 \to 4, \qquad 4 \to 0. \tag{5.17}$$

In the present investigation, the final link from $4 \to 0$ proves that each qubit acts as a control and target within the entangling blocks.

Let $S(\tau)$ denote the complete temporal-encoding operator, $V_\ell(\boldsymbol{\theta}_\ell)$ denote the trainable rotation block, and E denote the circular CNOT operator. The transformation associated with the layer ℓ can then be written as

$$U_\ell(\tau, \boldsymbol{\theta}_\ell) = E\, V_\ell(\boldsymbol{\theta}_\ell)\, S(\tau). \tag{5.18}$$

Consequently, the final quantum state is

$$|\psi(\tau, \boldsymbol{\theta})\rangle = U_3(\tau, \boldsymbol{\theta}_3) U_2(\tau, \boldsymbol{\theta}_2) U_1(\tau, \boldsymbol{\theta}_1) \, |00000\rangle \, . \tag{5.19}$$

As $S(\tau)$ is utilized in each layer, the input data will be fed repeatedly throughout the circuit instead of encoding upon entering the circuit. In this way, this approach improves the functionality of the measured temporal data through data re-uploading.

5.5. Pauli expectation measurement layer

For obtaining the quantum representation, we measure the expected values for Pauli-X, Pauli-Y, and Pauli-Z determined on each qubit, $P \in \{X, Y, Z\}$. For this, we define

$$z_q^P(\tau, \boldsymbol{\theta}) = \langle \psi(\tau, \boldsymbol{\theta})| \, P_q \, |\psi(\tau, \boldsymbol{\theta})\rangle \, , \tag{5.20}$$

where

$$P_q = I^{\otimes q} \otimes \sigma_P \otimes I^{\otimes(4-q)}, \tag{5.21}$$

is the five-qubit embedding of the selected Pauli operator.

The measured quantum feature vector is ordered as

$$\boldsymbol{z}^{(Q)}(\tau,\boldsymbol{\theta}) = \begin{bmatrix} z_0^X & \cdots & z_4^X & z_0^Y & \cdots & z_4^Y & z_0^Z & \cdots & z_4^Z \end{bmatrix}^{\mathrm{T}} \in [-1,1]^{15}. \tag{5.22}$$

These bound values can be derived directly from the eigenvalue set $\{-1, 1\}$ associated with every Pauli observable. These expectation values create a real-valued and continuous link between the quantum circuit and the classical regression network.

5.6. Classical regression head

The 15-dimensional quantum feature vector serves as an input to the classical fully connected neural network having two hidden layers with the width of 24. The normalized prediction becomes

$$\boldsymbol{h}^{(1)} = \tanh\left(W^{(1)}\boldsymbol{z}^{(Q)} + \boldsymbol{b}^{(1)}\right), \tag{5.23}$$

$$\boldsymbol{h}^{(2)} = \tanh\left(W^{(2)}\boldsymbol{h}^{(1)} + \boldsymbol{b}^{(2)}\right), \tag{5.24}$$

$$\widehat{\boldsymbol{y}} = \sigma\left(W^{(3)}\boldsymbol{h}^{(2)} + \boldsymbol{b}^{(3)}\right), \tag{5.25}$$

where $\tanh(\cdot)$ and $\sigma(\cdot)$ operate element-wise and

$$\sigma(v) = \frac{1}{1+\exp(-v)}. \tag{5.26}$$

The parameter dimensions are

$$W^{(1)} \in \mathbb{R}^{24\times 15}, \qquad \boldsymbol{b}^{(1)} \in \mathbb{R}^{24}, \tag{5.27}$$

$$W^{(2)} \in \mathbb{R}^{24\times 24}, \qquad \boldsymbol{b}^{(2)} \in \mathbb{R}^{24}, \tag{5.28}$$

$$W^{(3)} \in \mathbb{R}^{5\times 24}, \qquad \boldsymbol{b}^{(3)} \in \mathbb{R}^{5}. \tag{5.29}$$

The sigmoid output is consistent with the normalization introduced in Eq (5.11).

We utilize a regression head based on the mathematical model of the two hidden layer structure using the activation functions like tanh instead of reflects linear unit (ReLU). Both hidden layers use hyperbolic tangent activation functions, while the output uses the sigmoid function. The smooth bounded nonlinear activation function offered by tanh activation suits well-normalized quantum expectation features, whereas the sigmoid function confines predictions between $(0, 1)$, which makes the scaling $[0.05, 0.95]$ valid in this study. Even though ReLU can serve as another option with different properties of optimization and extrapolation, no ablation studies have been conducted on the activation functions in this work.

5.7. Trainable parameter count

Each of the three quantum layers contains three trainable rotations on each of the five qubits. The number of trainable quantum parameters is therefore

$$P_Q = 3nL = 3(5)(3) = 45. \tag{5.30}$$

The classical component contains

$$P_1 = 15(24) + 24 = 384, \tag{5.31}$$

$$P_2 = 24(24) + 24 = 600, \tag{5.32}$$

$$P_3 = 24(5) + 5 = 125, \tag{5.33}$$

and hence

$$P_C = P_1 + P_2 + P_3 = 1109. \tag{5.34}$$

The total number of jointly optimized parameters is consequently

$$P_{\text{total}} = P_Q + P_C = 45 + 1109 = 1154. \tag{5.35}$$

The architecture developed is based on a quantum classical surrogate where the temporal information is repeatedly encoded into a circuit comprising five qubits. Here, the rotations and circular CNOT block produce correlated states quantum mechanically. In addition, 15 Pauli expectation values represent the inputs for classical multi output regression head. All quantum calculations take place in a state-vector simulator. Hence, the numerical results indicate the approximation ability of the hybrid architecture implemented.

5.8. Training and evaluation protocol

The HQNN is trained against numerical reference trajectories generated on the temporal grid

$$t_n = \frac{nT}{N}, \qquad n = 0, \ldots, N, \tag{5.36}$$

with the terminal time $T = 100$ and $N = 2^{10} = 1024$ integration subintervals. The complete reference trajectory therefore contains $N+1 = 1025$ temporal observations. A subset of $N_{\text{tr}} = 129$ approximately equidistant indices is selected from this grid for supervised training. The scaling statistics (m_j, M_j) are computed from the complete reference trajectory and subsequently applied to the selected training targets.

For the normalized training times $\{\tau_i\}_{i=1}^{N_{\text{tr}}}$, the HQNN produces the batched prediction matrix

$$\widehat{\widetilde{Y}} = \begin{bmatrix} \widehat{\widetilde{\boldsymbol{y}}}(\tau_1)^{\mathrm{T}} \\ \vdots \\ \widehat{\widetilde{\boldsymbol{y}}}(\tau_{N_{\text{tr}}})^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{N_{\text{tr}} \times 5}. \tag{5.37}$$

Let $\widetilde{Y} \in \mathbb{R}^{N_{\text{tr}} \times 5}$ denote the corresponding normalized reference values. The data-fitting component of the objective is the mean squared error

$$\mathcal{L}_{\text{data}} = \frac{1}{5N_{\text{tr}}} \sum_{i=1}^{N_{\text{tr}}} \sum_{j=1}^{5} \left[\widehat{\widetilde{y}}_j(\tau_i) - \widetilde{y}_j(\tau_i)\right]^2. \tag{5.38}$$

Additional emphasis is placed on reproducing the initial state through the penalty

$$\mathcal{L}_{\text{IC}} = \frac{1}{5} \sum_{j=1}^{5} \left[\widehat{\widetilde{y}}_j(0) - \widetilde{y}_j(0)\right]^2. \tag{5.39}$$

The total training objective is

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda_{\text{IC}}\mathcal{L}_{\text{IC}}, \qquad \lambda_{\text{IC}} = 10. \tag{5.40}$$

The initial condition contribution is a supervised endpoint penalty; it is not a residual of the governing dynamic system. Consequently, its inclusion does not change the classification of the model as a supervised surrogate.

All computations are performed in double precision. The complete set of 1154 quantum and classical parameters is optimized jointly for 5000 epochs using the Adam algorithm with the initial learning rate

$$\eta_0 = 3 \times 10^{-3}. \tag{5.41}$$

A `ReduceLROnPlateau` scheduler reduces the learning rate by a factor of 0.5 when the total loss does not improve over 75 successive scheduler evaluations, subject to the lower bound

$$\eta_{\min} = 10^{-5}. \tag{5.42}$$

To control excessively large parameter updates, the global gradient norm is clipped at

$$\|\nabla_{\boldsymbol{\Theta}}\mathcal{L}\|_2 \leq 5. \tag{5.43}$$

A computational framework is built using the Python programming language with the help of several packages such as NumPy, PyTorch, PennyLane, and Matplotlib. For the purpose of integration, noise generation, normalization of data, and error calculations, we use NumPy. For designing, auto-differentiation, and optimization of both classical and learnable parameters in the models, PyTorch has been used. Five-qubit parametric quantum circuit has been designed using PennyLane on `default.qubit`. We perform the computation of quantum gradients using the backpropagation method on the simulator via PyTorch. Graphical plotting was done using Matplotlib. Other important packages `argparse`, `csv`, `json`, `pathlib`, and `random` from the standard library have been used for parameters configuration, storage of data, file manipulation, and random numbers control, respectively. We use a seed value of 100 to initialize the Python, NumPy, and PyTorch random number generators in the double-precision format.

Both deterministic and stochastic systems have been solved separately. For the deterministic problem, a reference path has been obtained setting the noise-intensity switching mechanism equal to zero. However, in the case of stochastic system, the noise-intensity switching mechanism is used as one. Moreover, the prescribed noise amplitude components are included in the reference solution for the stochastic case. Therefore, each trained neural-network learns only one particular reference trajectory.

After optimizing the network, the trained HQNN model is validated by evaluating the output through all 1025 time-steps in one forward pass in a batched manner. The normalized output is inverse scaled to physical units using Eq (5.12). The trajectory outputs are compared with the reference numerical solutions based on component wise errors and goodness of fit. Finally, we store the trained state dictionary, history loss, prediction plots, feature diagnosis, and performance measurements for post processing purposes.

5.9. Held-out temporal evaluation and validation protocol

The initial HQNN reconstruction was obtained using numerical observations selected over the complete time interval, and its reported full-grid errors measure the agreement with the reference solution. However, the same errors cannot be regarded as independent evidence of prediction accuracy at unobserved time points. For this reason, an additional temporal evaluation was carried out for the HQNN architecture ablations and the repeated stochastic realization experiments.

In the completed experiments, the *blocked interpolation* procedure was used to separate training, validation, and testing observations. Time positions were normalized to the interval $[0, 1]$. The observations in the interior interval $[0.40, 0.60)$ were reserved for testing, whereas the interval $[0.20, 0.30)$ was reserved for validation. The training candidates were selected from $[0, 0.20)$, $[0.30, 0.40)$, and $[0.60, 1]$. From these permitted regions, 129 approximately equally spaced temporal observations were selected for training. None of the validation or test observations was included in the training set. Since training observations occur on both sides of the withheld testing interval, this experiment examines *temporal interpolation*, not prediction over a future time window.

The five state variables have different numerical ranges. Therefore, for compartment j, the normalization constants were calculated only from the training observations. The normalized reference values are given by

$$\widetilde{y}_j(t) = 0.05 + 0.90\,\frac{y_j(t) - m_j^{\mathrm{tr}}}{d_j^{\mathrm{tr}}}, \qquad m_j^{\mathrm{tr}} = \min_{i \in \mathcal{I}_{\mathrm{tr}}} y_j(t_i), \qquad d_j^{\mathrm{tr}} = \max\left\{ \max_{i \in \mathcal{I}_{\mathrm{tr}}} y_j(t_i) - m_j^{\mathrm{tr}}, 10^{-12} \right\}. \quad (5.44)$$

The same training-derived constants were used during validation, testing, and inverse transformation of the predictions to the original model scale. Thus, the values belonging to the withheld interval were not used to determine the normalization parameters. Moreover, the NRMSE denominator was obtained from the corresponding training range rather than the test observations. This procedure differs from the original full-grid reconstruction analysis and helps avoid leakage of the testing information through the normalization step.

The completed HQNN was fitted with the Adam optimizer with an initial learning rate of 3×10^{-3}, a maximum of 5000 epochs, and double-precision arithmetic. The training objective combined the mean squared trajectory error with an initial condition penalty of weight 10. Gradients were clipped to a global norm of 5. A learning rate scheduler reduced the step size when validation loss stopped improving, and the model state with the lowest validation MSE was retained. Training was stopped early when the validation MSE did not improve for 750 consecutive epochs. The test observations were used only after selecting the trained model.

Each HQNN ablation was repeated with the network initialization seeds 100, 101, and 102 for the deterministic reference and for one selected stochastic reference path. Furthermore, the complete five-qubit HQNN was trained under the same three initializations for each of five independently generated stochastic paths. In these studies, the test MAE, RMSE, NRMSE, and R^2 were computed separately for the five model compartments. The results from different network initializations were distinguished from variation across independent stochastic realizations. It should be emphasized that training a separate HQNN for each stochastic path examines sample path reconstruction variability; it does not establish prediction of an unseen stochastic realization by a single trained network.

5.10. Ablation analysis of the HQNN architecture:

An architectural ablation study was carried out to assess which components of the proposed hybrid quantum classical surrogate contribute to held-out temporal accuracy. In particular, the investigation separates the effects of temporal data re-uploading, circular entanglement, the qubit count, and the measured observable set. The unmodified five-qubit, three-layer architecture with repeated input encoding, a circular CNOT ring, and Pauli-X, Pauli-Y, and Pauli-Z expectations is used as the reference configuration. Its quantum circuit has 45 trainable rotation angles, yielding 15 expectation features for the 15-24-24-5 classical regression head and 1154 trainable parameters in total.

In all variants, the classical hidden widths remain 24 and 24, and the final five outputs correspond to X, B, C, R, and E. Differences in the total parameter count introduced by changes in features' dimension are reported explicitly.

For n qubits, $L = 3$ trainable layers, and d measured expectation features, the resulting parameter count for this fixed-width classical head is

$$P_{\mathrm{abl}}(n, d) = 3nL + (24d + 24) + (24^2 + 24) + (24 \cdot 5 + 5) = 9n + 24d + 749. \tag{5.45}$$

The comparison configurations and corresponding parameter counts are listed in Table 4.

Table 4. Pre-specified HQNN architecture ablations. All variants have three quantum layers and classical hidden widths of 24 and 24.

Variant	Qubits	Features	Parameters
Full HQNN (reference)	5	15	1154
No data re-uploading	5	15	1154
No CNOT entanglement	5	15	1154
Three-qubit HQNN	3	9	992
Four-qubit HQNN	4	12	1073
Six-qubit HQNN	6	18	1235
Pauli-Z measurements only	5	5	914
Pauli-X and Pauli-Z only	5	10	1034

Each variant was trained independently with the same training-sample, initial condition penalty, optimizer settings, temporal partitions, and training-only normalization specified in Section 5.9. The fixed-width comparison makes it possible to examine changes in the quantum circuit's features. However, because the input features dimension varies among configurations, the total number of trainable parameters is not identical for all variants. Accordingly, these ablations do not establish a parameter-count-independent architectural advantage. For the completed stochastic ablation, all eight configurations were fitted to the same Brownian reference path (seed=100). Separate Brownian paths were not included in that experiment; variability across optimization seeds is not interpreted as between-realization uncertainty.

Performance was assessed separately for each biological compartment, particularly the cardiotoxicity compartment $E(t)$, using held-out MAE, RMSE, NRMSE, and R^2, as well as the training wall-clock time and the number of learnable parameters. To quantify the difference from the full HQNN, let $e^{(v)}_{s,a,j}$ denote the held-out NRMSE for variant v on noise realization s, initialization a,

and compartment j. The paired error difference is

$$\Delta^{(v)}_{s,a,j} = e^{(v)}_{s,a,j} - e^{(\text{full})}_{s,a,j}. \tag{5.46}$$

6. Results and discussion

In the deterministic system, the HQNN reconstructs a smooth numerical reference trajectory. For the stochastic system, it reconstructs a selected simulated stochastic path. Therefore, the stochastic results should be interpreted as an approximate reconstruction of a single realizable stochastic trajectory rather than as a deterministic representation of the ensemble behavior of the stochastic system.

6.1. Deterministic system reconstruction analysis

Figure 2 describe the quantum component of the surrogate. The process is started by encoding the temporal scalar inputs inside the five-qubit register. Then trainable rotation gates along with entanglement gates act upon the encoded states. The classical regression network receives only the outputs generated from the measurement process, which means that the amplitude information of the quantum states is not fed directly into the classical network. Although there is a correspondence between the number of qubits and biological compartments, no direct mapping exists between them. In other words, all qubit information acts collectively towards predicting a biological compartment.

Figure 3 presents the available numerical reference trajectories for the deterministic five-compartment breast cancer model. The numerical solution is computed over the time interval $0 < t < 100$. The trajectories are considered as references to examine the reconstruction ability of the HQNN model. The numerical parameters and initial states used in the code are listed in Table 2.

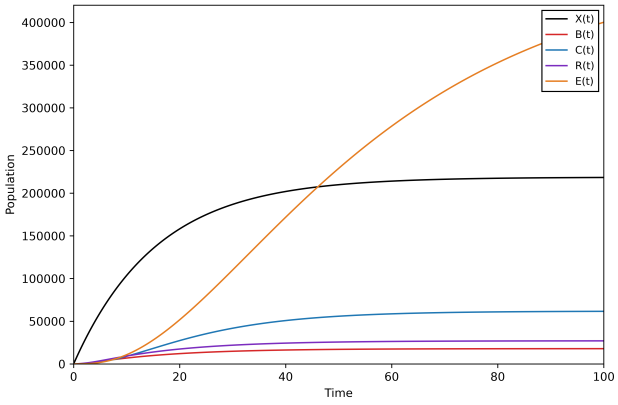

Figure 3. Deterministic reference trajectories for the five model compartments over $0 \leq t \leq 100$.

Figure 4 shows the results of the training loss function that were analyzed for the HQNN model. The training loss declines very sharply in the beginning, followed by slow decline, obtaining the loss magnitude of the order of 10^{-5}. It can be seen that there was good convergence achieved in the training phase. The rapid decline in the loss shows that the network could capture major features in the beginning phases, while a slow decline means an improvement in the approximating ability of the

network. The loss function graph cannot confirm the prediction or reconstruction capability of the model.

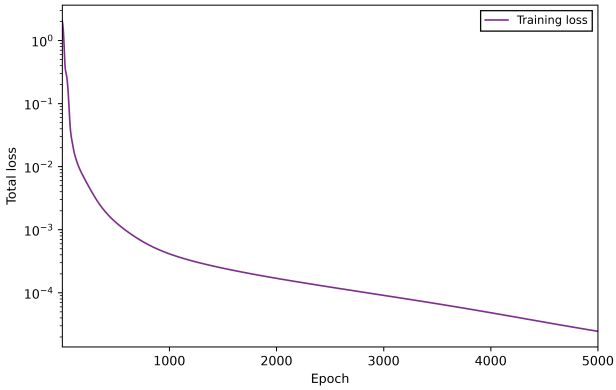

Figure 4. Training loss over 5000 iterations for the original deterministic HQNN reconstruction.

Figure 5 shows the comparison between the numerical reference solutions and the corresponding HQNN reconstructions. The reconstructed trajectories exhibit close agreement with the reference solutions over the considered time interval. In particular, the HQNN successfully reproduces the dominant temporal behavior of all five state variables, despite their different physical scales and dynamic ranges. The close correspondence between the two sets of trajectories indicates that the learned representation is capable of approximating the nonlinear temporal dynamics encoded in the reference solution.

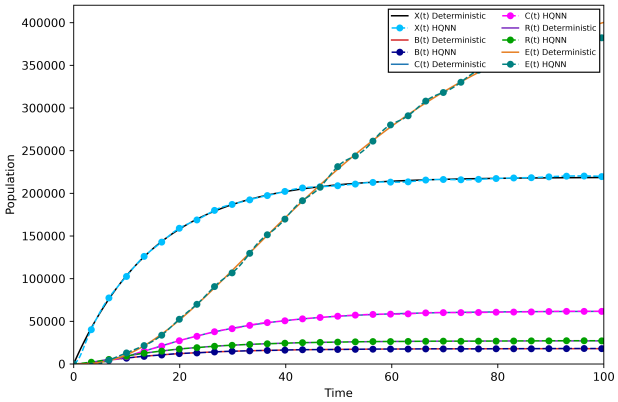

Figure 5. Original full-grid deterministic trajectory comparison between the numerical reference and the corresponding hybrid network reconstruction for X, B, C, R, and E.

The dimensional errors are presented in Figure 6. It is noted that there is higher inconsistency within the state variables E and X than the other variables. This is because that there exist large absolute dimension among the two variables. Thus, dimensional inconsistency must be understood in relation to the relative performance measures and not considered separately. In this regard, the determination coefficient R^2 provides an independent measure to evaluate accuracy. From our analysis, it has been found that there is very close proximity toward one for all five state variables, ranging between 0.999367 and 0.999836.

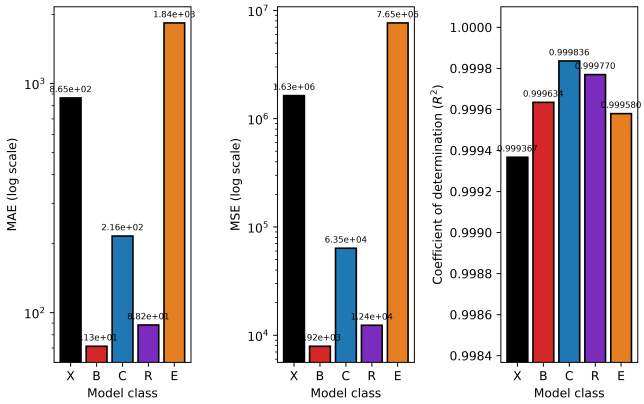

Figure 6. Compartment-wise mean absolute error, mean squared error, and coefficient of determination for the original deterministic full-grid reconstruction.

All compartments of the model and the NMAE are shown in Figure 7. The corresponding NMAE values are 3.96×10^{-3}, 4.00×10^{-3}, 3.51×10^{-3}, 3.27×10^{-3}, and 4.59×10^{-3} for X, B, C, R, and E, respectively. Thus, all NMAE values remain below 0.46% of the corresponding reference range. Figure 8 depicts the NRMSE across all compartments of the model. The corresponding NRMSE values are 5.85×10^{-3}, 4.98×10^{-3}, 4.09×10^{-3}, 4.12×10^{-3}, and 6.91×10^{-3} for X, B, C, R, and E, respectively. Therefore, we conclude that the NRMSE values for the deterministic approach are less than 0.7% for all the simulations. Furthermore, we note the consistency between the NMAE and NRMSE results; therefore, we can conclude that there are no prominent localized errors. A separate NMSE figure is unnecessary because, under the adopted normalization, NMSE = NRMSE^2; consequently, the NMSE can be directly inferred from the reported NRMSE values without introducing redundant graphical information.

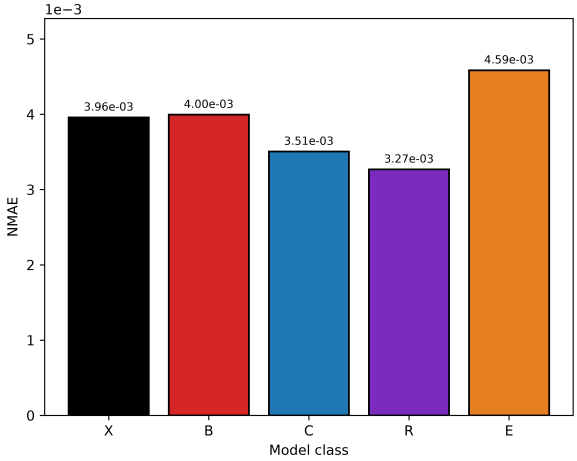

Figure 7. Deterministic normalized mean absolute errors for the five states.

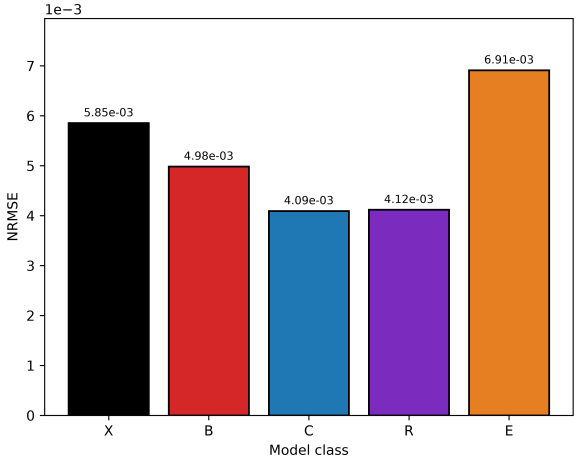

Figure 8. Deterministic normalized root mean squared errors for the five states.

Figure 9 shows a mathematical model based on the HQNN that produces 15 different channels. These channel results show boundedness and dependence on time. This observation implies that there exists a mapping of the single scalar temporal input through the PQC to generate these nonlinear feature representations. In other words, this diverse group of features provides several forms of temporal input representation that help improve the approximation capability toward modeling the nonlinear trajectory form. However, it should be noted that simply obtaining these feature representations cannot guarantee that they are better than classical features; otherwise, a comparison test must be conducted.

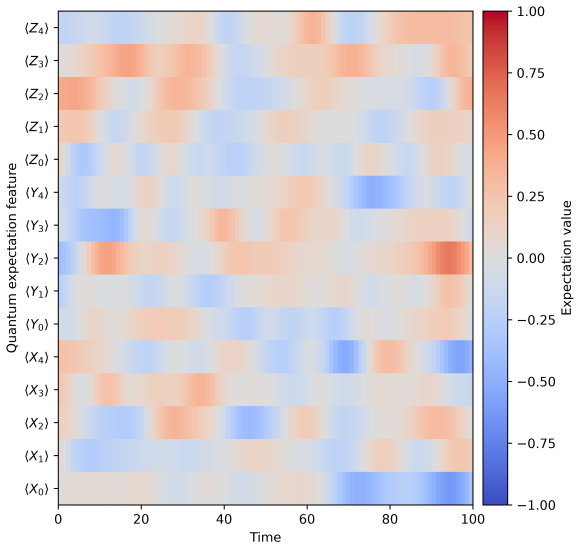

Figure 9. Heatmap of the 15 deterministic Pauli expectation features. Rows denote the measured qubit–observable channels, columns denote time, and the color scale spans $[-1, 1]$.

Table 5 summarizes the deterministic reconstruction criteria of the HQNN for all five compartments. High values of R^2 and low values of NMAE and NRMSE indicate the precision of the approximation of the HQNN compared with the reported reference solution over the interval $0 < t < 100$. The deterministic reconstruction is considered as a robust benchmark to evaluate the

efficiency of the HQNN against the stochastic disturbance.

Table 5. Deterministic HQNN reconstruction metrics.

State	MAE	RMSE	NMAE	NRMSE	R^2
X	8.65×10^2	1.28×10^3	3.96×10^{-3}	5.85×10^{-3}	0.999367
B	6.13×10^1	8.89×10^1	4.00×10^{-3}	4.98×10^{-3}	0.999634
C	2.16×10^2	2.52×10^2	3.51×10^{-3}	4.09×10^{-3}	0.999836
R	8.82×10^1	1.11×10^2	3.27×10^{-3}	4.12×10^{-3}	0.999770
E	1.84×10^3	2.77×10^3	4.59×10^{-3}	6.91×10^{-3}	0.999580

6.2. Results of the stochastic case

The stochastic cancer cardiotoxicity numerical reference trajectory is depicted in Figure 10. It is shown that in contrast to deterministic cases, there is some local variation due to the increment of randomness within the considered model. This variation is due to the random perturbations used in the stochastic model. This results in sample paths that are usually rougher compared with those of deterministic models. The reference path shown in Figure 1 can therefore be taken as a particular instance of the stochastic model.

Figure 11 depicts the training loss of the HQNN during the process of end-to-end learning. The loss has a general trend of decreasing and reaches a relatively low level. This shows that the training successfully minimizes the difference between the output of the HQNN and the chosen stochastic reference path. There is some residual loss after reaching the low plateau level. This is quite natural. In particular, the HQNN learns the sample path by means of a smooth parameterization while the stochastic reference solution has some random increment in its localized part. Thus, the residual loss can be considered as a quantification of the error when approximating a non smooth stochastic sample path by a low dimensional smooth model.

Figure 12 shows a comparison between the stochastic reference trajectory and the reconstructed stochastic trajectory from the HQNN approach. In the proposed approach, it is shown that the HQNN successfully recovers the major trends, increases, decreases, and transitions of the individual state variables without considering the strongest stochastic variations. The behavior of the HQNN matches its objective as the replacement reconstruction technique because the major structure of the specific realization is recovered without replicating all individual stochastic increments. Hence, the compeers sample path reconstruction shows that the HQNN has accurately recovered the sample path, and no stochastic learning has occurred through this process.

The numerical outcomes presented in Figure 13 give an indication of the degree of agreement between the HQNN's prediction and the selected stochastic path. All those measures must be considered according to the context of sample path approximation. It must be noted here that higher R^2 indicates a better agreement with respect to the changes observed in the selected sample path; however, it does not prove that the HQNN has successfully learned the drift and diffusion parameters of the stochastic differential equation system. This study has not been performed for different stochastic sample paths produced via distinct seeds.

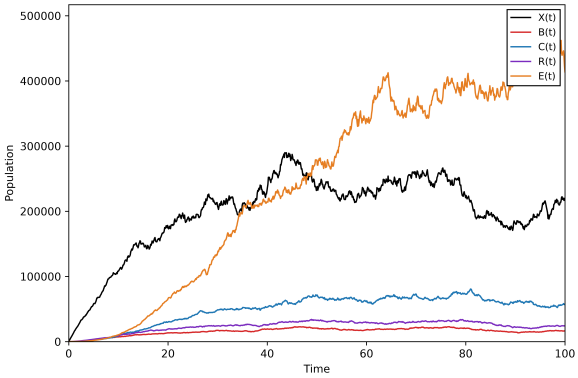

Figure 10. Stochastic numerical reference trajectory for the five cancer cardiotoxicity compartments.

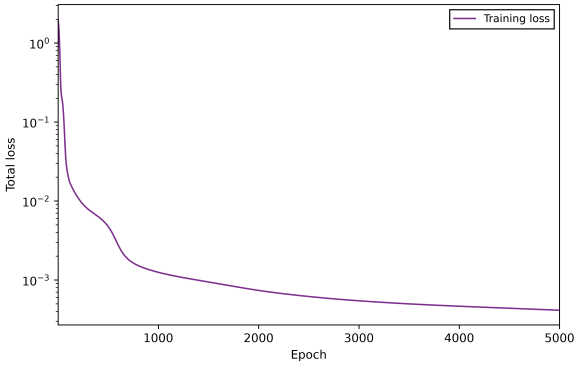

Figure 11. Training loss of the stochastic HQNN fitted to the selected sample path.

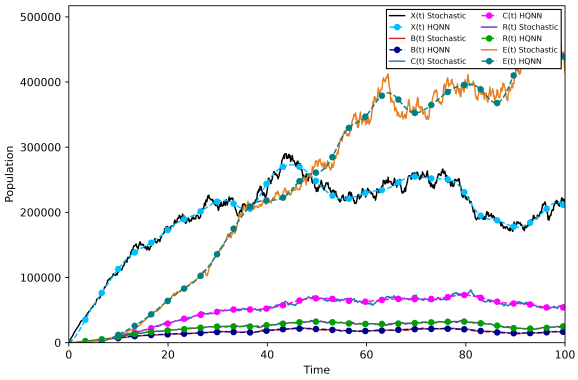

Figure 12. Numerical reference and hybrid network reconstruction for one selected stochastic sample path in the five compartments.

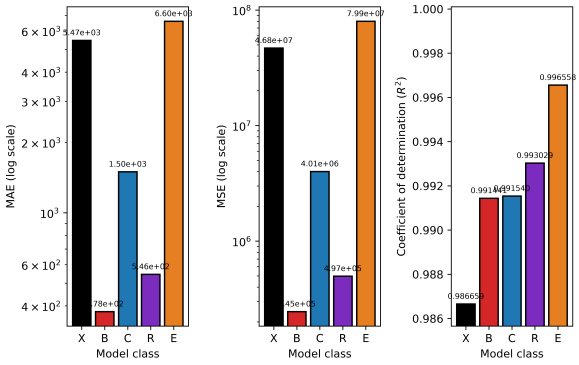

Figure 13. Stochastic sample path MAE, MSE, and R^2 for the five compartments.

Figure 14 presents the NMAE is presented for different state variables in the model. Normalizing helps compare the performance of compartments because of differences in the scales and ranges caused by the normalization procedure. Therefore, NMAE provides the relative performance of the reconstructed compartments compared to other compartments. In addition, the results reveal the challenging compartments for the HQNN approximation in terms of the selected stochastic paths.

Because the NRMSE assigns greater weight to larger deviations than the NMAE as shown in Figure 15, the NRMSE offers additional insights into the reconstruction accuracy. High NRMSE values suggest deviations due to sharp stochastic jumps within the trajectory. Jointly considering both metrics, NMAE and NRMSE, gives more informative results compared with using each measure alone, particularly for trajectories with irregular fluctuations.

Figures 16 and 17 we analyze the temporal evolution of the features channels obtained from the stochastic HQNN. It produces several bound and non-constant temporal channels. This indicates that the PQC converts the time input scalar value into a complex nonlinear feature representation. These channels must be understood as computational latent variables utilized by the network to construct an output approximation. These channels do not represent biological processes or the significance of state-variables independently.

Figure 18 demonstrates the learned 45 angles in the quantum rotations during the training process of the entire network. The variability shown demonstrates the role played by the quantum rotation parameters during the optimization process of the reconstruction problem. However, the value of each individual angle cannot be considered a typical way of estimating the importance of features in a traditional sense. Due to the periodic behavior of quantum rotation parameters, various parameter sets give rise to similar quantum computations. Hence, parameter magnitudes alone are not sufficient for evaluating the importance of the individual parameter or input feature.

Table 6 summarizes the reconstruction performance of the HQNN for the selected stochastic sample path. The NMAE values range from 1.34×10^{-2} to 1.89×10^{-2}, while the corresponding NRMSE values range from 1.82×10^{-2} to 2.48×10^{-2}. We find the range of R-squared values varying from 0.986659 to 0.996558. These values show good agreement between the reconstructed trajectory and the reference trajectory. This proves that the HQNN learns the major temporal structure of the stochastic realization under consideration. The higher errors than those observed in the deterministic process can be attributed to the added difficulty involved in reconstructing local stochastic perturbations.

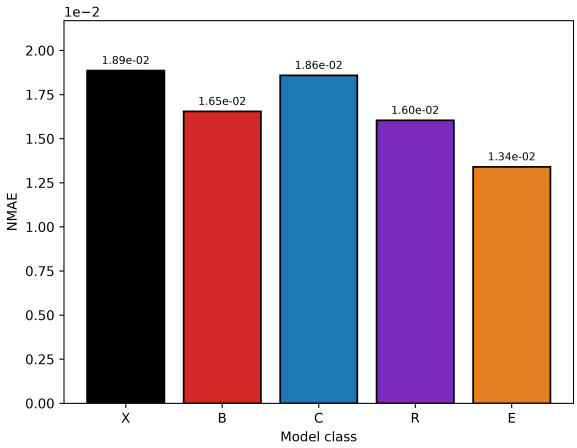

Figure 14. Stochastic range normalized mean absolute errors for the selected sample path.

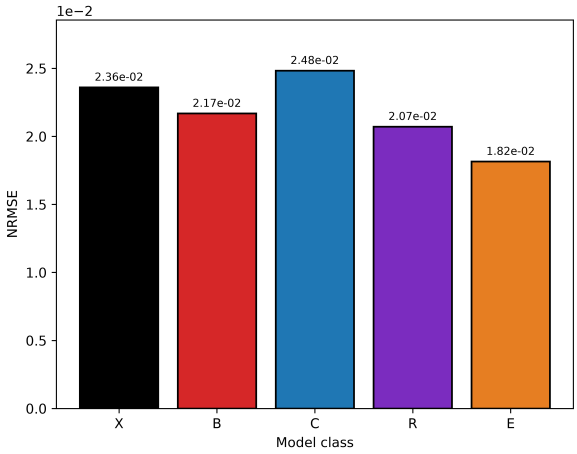

Figure 15. Stochastic range normalized root mean squared errors for the selected sample path.

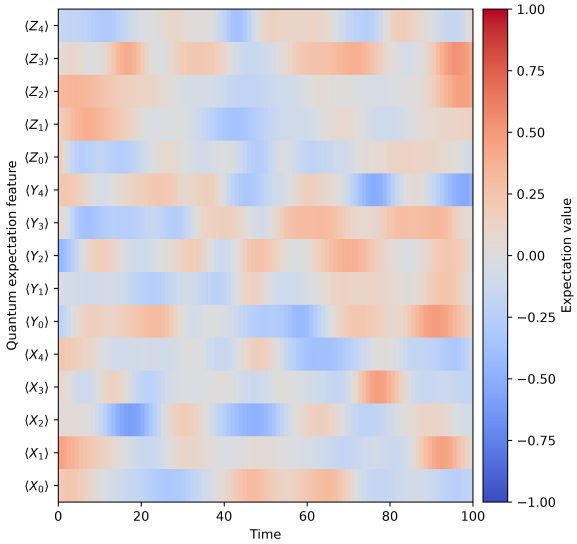

Figure 16. Heatmap of the 15 trained Pauli expectation features for the stochastic HQNN.

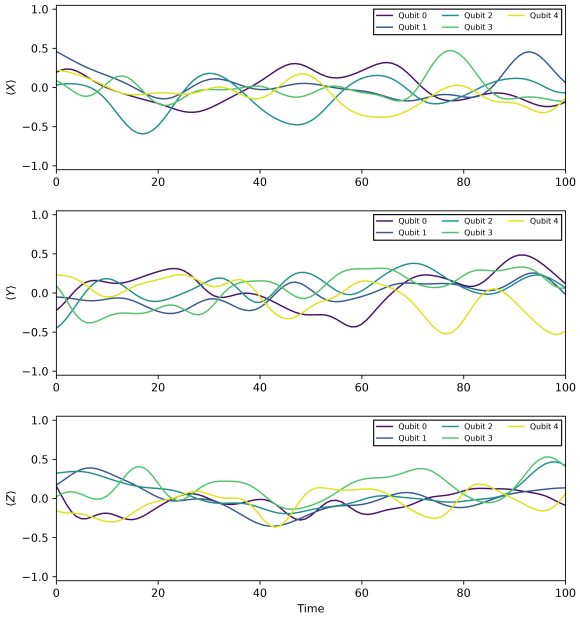

Figure 17. Pauli-X, Pauli-Y, and Pauli-Z expectation trajectories of the trained stochastic circuit.

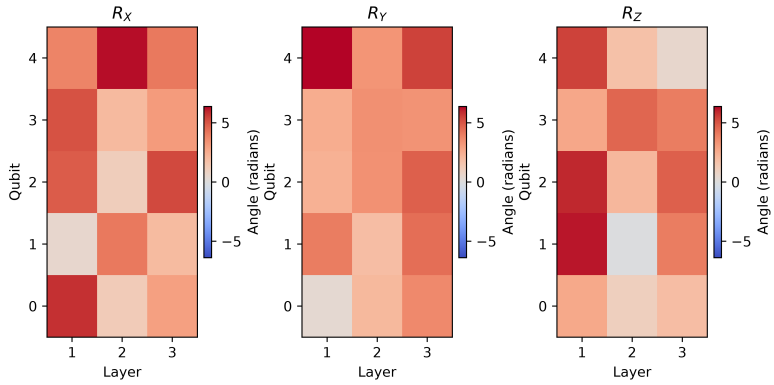

Figure 18. Optimized stochastic HQNN rotation angles by layer, qubit, and rotation axis.

Table 6. Stochastic HQNN sample-path reconstruction metrics.

State	MAE	RMSE	NMAE	NRMSE	R^2
X	5.47×10^3	6.84×10^3	1.89×10^{-2}	2.36×10^{-2}	0.986659
B	3.78×10^2	3.81×10^2	1.65×10^{-2}	2.17×10^{-2}	0.991449
C	1.50×10^3	2.00×10^3	1.86×10^{-2}	2.48×10^{-2}	0.991540
R	5.46×10^2	7.05×10^2	1.60×10^{-2}	2.07×10^{-2}	0.993029
E	6.60×10^3	8.94×10^3	1.34×10^{-2}	1.82×10^{-2}	0.996558

Overall, the observations show the successful performance of sample path reconstruction by the HQNN; however, these results apply only to the considered stochastic realization and cannot provide

evidence for generalization to new independent stochastic realizations or learning the full stochastic process dynamics of the considered stochastic cancer cardiotoxicity system.

The separate held-out ablation results in Section 6.4 identify substantial out-of-sample limitations. The additional five-path experiment in Section 6.3 quantifies the variability across separately trained stochastic paths, but does not establish transfer to previously unseen paths by one trained network. Classical simulation utilizing an exact classical state vector simulator cannot prove the finite shots effect and hardware noise effects on the simulations. Moreover, the lack of a classical benchmark with parameter matching cannot prove any advantage in quantum simulations.

6.3. Validation on five independent stochastic trajectories

The stochastic reconstruction reported earlier in this section considers a particular reference realization. However, the behaviour of the HQNN may change when the stochastic increments and the initial values of the neural network weights are changed. To explore this matter, an additional investigation was performed on five independently generated stochastic reference trajectories. The reference paths were indexed by the Brownian seeds 100, 101, 102, 103, and 104. For each path, the full five-qubit HQNN was trained three times with the network initialization seeds 100, 101, and 102. Therefore, a total of 15 HQNN fits and 75 compartment-level test metric records were obtained. Each fit was assessed on a held-out blocked temporal interval, with training-only normalization and early stopping based on a separate validation interval.

The calculation of statistical uncertainty was performed in two stages. First, the three network initialization results were averaged within each stochastic realization. Next, these five path-level averages were used to calculate the mean error, sample standard deviation and Student's-t 95% confidence interval. The five independent paths, rather than the fifteen network fits, are the statistical units for the confidence intervals. The resulting path level values were used to form the Student-t confidence intervals reported below. This analysis measures the reproducibility of separately fitted sample path reconstructions and not the ability of one trained network to predict an unseen stochastic realization.

Table 7 shows the values of NRMSE for the five state variables. The smallest mean NRMSE of 0.1098 was obtained for the cardiotoxicity compartment E, while the recovery compartment R had the largest mean NRMSE of 0.2816. The overall macro-NRMSE was 0.2238, with a 95% confidence interval from 0.1614 to 0.2861. This interval was estimated from only five stochastic realizations and should be considered to be preliminary. The other compartments X, B, and C also showed differences in their reconstruction errors between the paths.

Moreover, the coefficient of determination in Table 8 was negative for every compartment. In particular, the across-path mean R^2 was -46.47 for R, -30.39 for B, -28.12 for C, -27.57 for X, and -0.66 for E. The overall macro mean was -26.64, with a 95% confidence interval of $[-39.50, -13.79]$. These negative values indicate that the fitted models did not reproduce the withheld temporal portions better than the corresponding test-subset mean predictor in squared-error terms. It is possible for the NRMSE to look moderate and for R^2 to be strongly negative when the withheld observations have comparatively small variation. Therefore, the two performance measures should be interpreted together.

Table 7. Mean ± sample standard deviation and 95% confidence intervals across five independent stochastic paths for held-out NRMSE and R^2. Three training initializations were averaged within each path.

State	Mean NRMSE ± SD	NRMSE 95% CI	Mean R^2
X	0.2512 ± 0.0818	[0.1496, 0.3528]	-27.57
B	0.2489 ± 0.0668	[0.1661, 0.3318]	-30.39
C	0.2273 ± 0.0464	[0.1697, 0.2850]	-28.12
R	0.2816 ± 0.0887	[0.1715, 0.3917]	-46.47
E	0.1098 ± 0.0237	[0.0804, 0.1392]	-0.66
Macro average	0.2238 ± 0.0502	[0.1614, 0.2861]	-26.64

Table 8. Cross-path mean R^2, sample standard deviation, and 95% confidence intervals from the same 15 full-HQNN fits. Negative R^2 represents worse squared-error performance than the mean of the withheld test subset.

State	Mean R^2 ± SD	R^2 95% CI
X	-27.575 ± 17.512	[-49.319, -5.831]
B	-30.393 ± 22.262	[-58.034, -2.751]
C	-28.120 ± 8.975	[-39.264, -16.976]
R	-46.474 ± 8.777	[-57.372, -35.575]
E	-0.658 ± 0.258	[-0.978, -0.337]
Macro average	-26.64 ± 10.36	[-39.50, -13.79]

Table 9. Dimensional MAE and RMSE across five stochastic reference trajectories. Units are those of the modelled compartment populations, and the values are not patient-specific clinical measurements.

State	MAE mean ± SD	MAE 95% CI	RMSE mean ± SD	RMSE 95% CI
X	58848±19152	[35068,82628]	62115±19499	[37905,86326]
B	4914±1432	[3136,6692]	5132±1465	[3313,6951]
C	15548±3944	[10651,20445]	16354±3896	[11516,21192]
R	8117±2516	[4993,11241]	8373±2567	[5186,11560]
E	42303±13602	[25415,59192]	50422±14369	[32581,68264]

Figures 19–23 illustrate the variation in the performance measurements. The MAE and RMSE plots show relatively large dimensional deviations for X and E, which is partly related to the larger magnitudes of their modeled population trajectories. Therefore, these dimensional measures should not be used alone to compare compartments. The normalized values in Figure 19 are more suitable for comparing the relative errors among the five model classes. Figure 23 shows that the overall normalized error also varies from one stochastic reference path to another. In the original experiment, the HQNN followed the dominant behaviour of one selected stochastic reference trajectory.

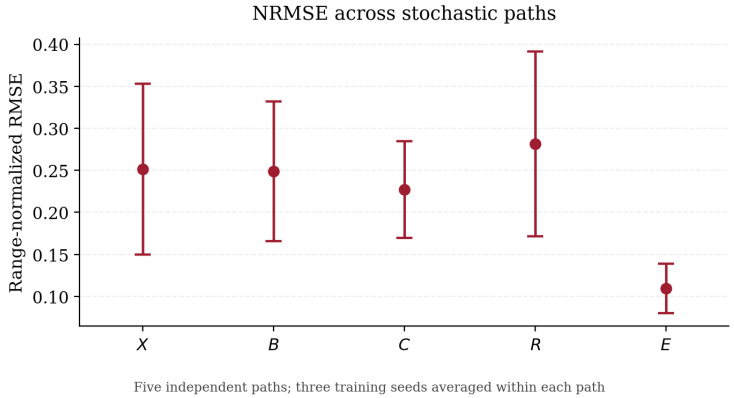

Figure 19. Compartment-wise mean NRMSE and 95% confidence intervals across five independent stochastic paths. Three HQNN initialization results were averaged within each path.

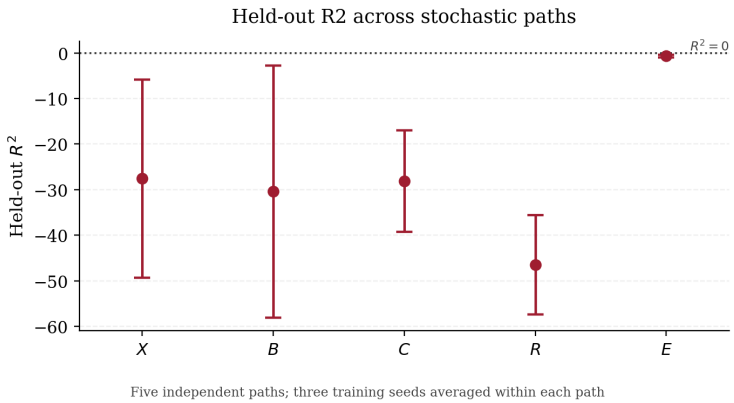

Figure 20. Mean held-out R^2 values with 95% confidence intervals across independent stochastic realizations. The horizontal dotted line marks $R^2 = 0$.

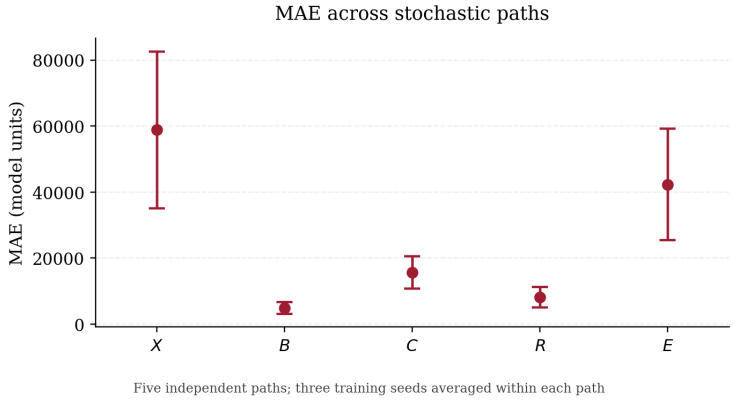

Figure 21. Mean absolute errors in model population units with 95% confidence intervals over five independently fitted stochastic trajectories.

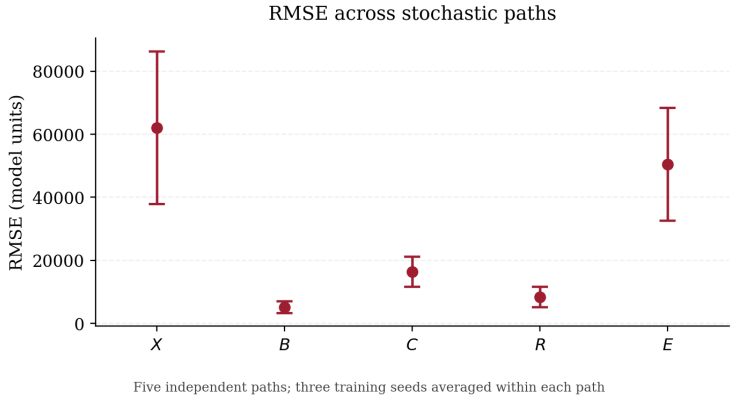

Figure 22. RMSEs in model population units with 95% confidence intervals over five independently fitted stochastic trajectories.

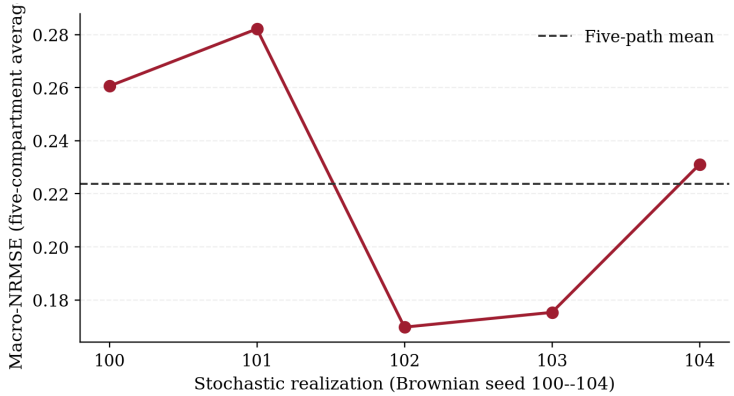

Figure 23. Variation in macro-NRMSE across the five stochastic reference paths. Each point is the average of three separate HQNN initializations.

6.4. Ablation analysis of the HQNN

To understand the contribution of the different components of the quantum circuit, an ablation investigation was performed for eight HQNN architectures. In addition to the full five-qubit circuit, we considered variants without temporal data re-uploading and without CNOT entanglement. The number of qubits was also changed to three, four and six, while the measured quantum features were restricted to Pauli-Z or Pauli-X/Z in two further variants. Each architecture was trained with three different initializations on the deterministic reference and three initializations on one selected stochastic reference, giving a total of 48 completed fits. The errors were calculated on held-out blocked temporal observations, and the results are reported in Table 10.

For a given initialization, the NRMSE values were first averaged over the five compartments X, B, C, R, and E to obtain the macro-NRMSE. Next, the mean and sample standard deviation were calculated over the three network initializations. In this way, the repeated network training was used to measure initialization variability without counting the five biological compartments as independent experimental replicates.

From Table 10, it can be seen that the lowest mean error in the deterministic case was obtained

when temporal data re-uploading was removed. The corresponding mean macro-NRMSE was 0.0512, whereas the full HQNN gave 0.1720. The model without entanglement also produced a smaller mean error of 0.0604. A similar comparison was made for the selected stochastic reference trajectory. In this case, the smallest mean macro-NRMSE was 0.1989 for the model without entanglement, while the full network gave 0.2606 and the model without re-uploading gave 0.2237. Therefore, the additional encoding and entangling operations were not necessary for achieving the lowest test errors in these particular experiments.

The ablation study was carried out using one deterministic trajectory and one selected stochastic trajectory. The three repetitions within each architecture refer to the neural network's initialization rather than three independent stochastic paths. For this reason, the ablation results are discussed separately from the five-path investigation in Section 6.3. The reference files in this ablation were supplied by the numerical data-generation code and are not independent evidence that the stochastic paths were obtained using the LSCM formulation. The present findings indicate empirical architecture sensitivity and do not establish a quantum computational advantage over classical methods.

Table 10. Ablation performance on held-out blocked-interpolation test data. Values are the mean ± sample standard deviation (SD) of macro-NRMSE over three training initialization seeds; a lower NRMSE is better. The stochastic column uses one stochastic reference trajectory, not an ensemble of stochastic realizations.

Variant	Parameters	Deterministic NRMSE	Stochastic NRMSE
Full HQNN (5 qubits)	1154	0.1720 ± 0.1814	0.2606 ± 0.1229
No temporal re-uploading	1154	0.0512 ± 0.0249	0.2237 ± 0.1263
No entanglement	1154	0.0604 ± 0.0319	0.1989 ± 0.0899
3 qubits	992	0.2030 ± 0.1644	0.2787 ± 0.1238
4 qubits	1073	0.1853 ± 0.1609	0.2520 ± 0.1345
6 qubits	1235	0.3419 ± 0.1011	0.3639 ± 0.1074
Pauli-Z only	914	0.3629 ± 0.0741	0.3673 ± 0.0220
Pauli-X/Z	1034	0.2430 ± 0.1721	0.2435 ± 0.1227

Table 11. Mean macro-R^2 on the held-out test subset and count of runs whose lowest validation loss occurred at Epoch 1 (out of three initialization seeds per case). Negative R^2 indicates worse squared-error performance than the test subset's mean value predictor.

Variant	Det. R^2	Stoch. R^2	Det. epoch-1	Stoch. epoch-1
Full HQNN (5 qubits)	−137.88	−21.31	1/3	2/3
No temporal re-uploading	−2.77	−15.55	0/3	0/3
No entanglement	−7.93	−12.19	0/3	0/3
3 qubits	−151.45	−25.01	1/3	2/3
4 qubits	−139.88	−19.45	1/3	2/3
6 qubits	−231.85	−37.67	0/3	1/3
Pauli-Z only	−360.60	−38.97	2/3	3/3
Pauli-X/Z	−176.52	−19.20	1/3	2/3

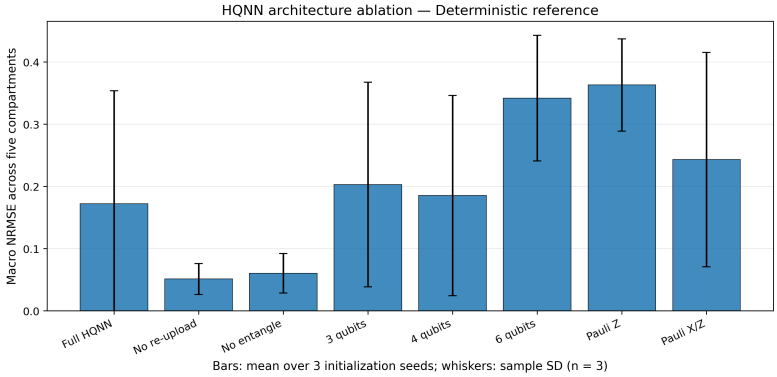

Figure 24. Mean macro-NRMSE over three initialization seeds for the eight HQNN variants, using the deterministic reference trajectory. Error bars represent sample standard deviation.

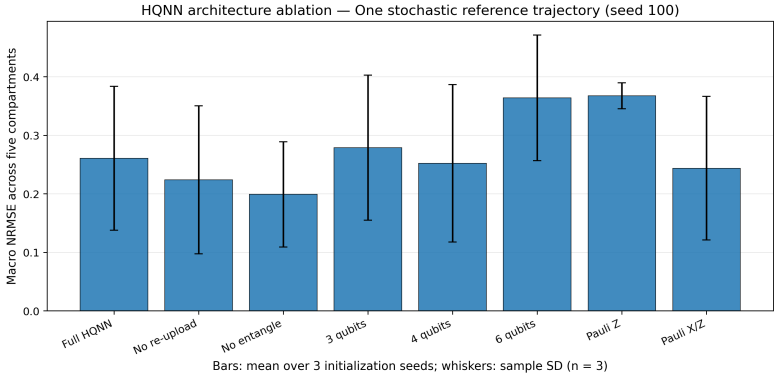

Figure 25. Mean macro-NRMSE over three initialization seeds for the eight HQNN variants, evaluated against one stochastic reference trajectory (seed=100). Error bars measure variation across initialization seeds, not across independent stochastic paths.

6.5. Biological interpretation and limitations

The mathematical compartments represent Stage I/II cancer X, Stage III cancer B, Stage IV cancer C, post-chemotherapy recovery R, and cardiotoxicity E. The model may be used to discuss the possible consequences of assumed cancer progression, recovery and cardiotoxicity-related rates. For example, the variable $E(t)$ provides a model-based representation of the number of patients experiencing a cardiotoxic complication under the prescribed transition coefficients. However, the present computations are not based on independent clinical patient follow-up data. Consequently, the trajectories cannot be used to estimate the treatment risk for any particular patient, nor can the network results be treated as clinical guidance.

The additional stochastic calculations show that the HQNN behaves differently across independent simulated trajectories, with unfavourable held-out R^2 values despite the original favourable full-grid reconstruction. Moreover, model parameter calibration, uncertainty ranges, Latin hypercube sensitivity or partial rank correlation analysis, unseen initial condition test,s and independent clinical validation have not been established in this work. All quantum circuit evaluations were performed using a classical state-vector simulator; finite-shot noise, hardware limitations, and quantum speedup

have not been tested.

7. Conclusions

In the present study, the deterministic and stochastic trajectories of an existing five-compartments breast cancer cardiotoxicity mathematical model were investigated using an HQNN. The proposed network contains five qubits and three quantum layers, in which temporal encoding, trainable rotations, and CNOT gates are used to obtain 15 Pauli expectation features. A classical regression head with two hidden layers was then applied to reconstruct the five model compartments simultaneously. The original full-trajectory numerical results showed close agreement between the selected reference curves and their HQNN reconstructions.

To explore the role of quantum circuit components, 48 ablation experiments were completed. It was observed that the network without temporal re-uploading gave the smallest mean held-out macro-NRMSE in the deterministic case, while the circuit without entanglement gave the lowest error on the selected stochastic realization. However, the full HQNN was not the best-performing architecture, and a large proportion of the held-out R^2 values were negative. Moreover, five independent stochastic trajectories were examined with three different HQNN initializations per path. The mean macro-NRMSE was 0.2238, with a 95% confidence interval of $[0.1614, 0.2861]$, whereas the corresponding mean held-out R^2 was -26.64. These findings indicate that the successful approximation of one numerical trajectory cannot be taken as evidence of reliable performance on new temporal observations or independent stochastic paths.

The current work therefore provides an investigation of the numerical behaviour and limitations of a simulated quantum classical surrogate. The stochastic simulations describe model-based variability and not measured patient uncertainty. In future studies, the numerical reference procedure should be reconciled with the theoretical numerical formulation. The supplementary mesh-refinement diagnostic supports the consistency of the implemented LSCM solvers. Since the circuit is evaluated through a classical state-vector simulator, no quantum computational advantage or direct clinical decision-making capability is concluded from the present results.

From a public health perspective, the main practical implication is the need for careful validation before numerical projections are considered in planning or monitoring activities. Health authorities and clinical research groups may prioritize the collection of longitudinal information on cancer stage, treatment response, recovery, recurrence and cardiovascular complications. Such data would allow the assumed model coefficients to be estimated and the resulting trajectories to be compared with observed outcomes. At its current stage, the proposed computational approach is suitable for methodological investigation and illustrative scenario analysis.

Author contributions

Azizah Alrajhi: Conceptualization, Writing–original draft, Formal analysis, Validation, Funding acquisition; Misbah Ullah: Conceptualization, Writing–original draft, Validation. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors gratefully acknowledge the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, for funding this project under grant No. (IPP: 1804-247-2026) and for providing technical support.

Conflict of interest

The authors declare no conflicts of interest.

References

1. I. Jahan, M. E. Chowdhury, S. Vranic, R. M. Al Saady, S. Kabir, Z. H. Pranto, et al., Deep learning and vision transformers-based framework for breast cancer and subtype identification, *Neural Comput. Appl.*, **37** (2025), 9311–9330. https://doi.org/10.1007/s00521-025-10984-2
2. O. Freihat, D. Sipos, A. Kovacs, Global burden and projections of breast cancer incidence and mortality to 2050: A comprehensive analysis of GLOBOCAN data, *Front. Public Health*, **13** (2025), 1622954. https://doi.org/10.3389/fpubh.2025.1622954
3. H. Gu, R. Wang, N. M. Beeraka, P. Lu, X. Zhao, L. Li, et al., Global burden and trends of breast cancer: GLOBOCAN 2022 estimates of incidence and mortality in 185 countries, *Chin. Med. J.*, **139** (2026), 404–414. https://doi.org/10.1097/CM9.0000000000003921
4. Y. Chen, R. Wang, L. Wu, J. Li, M. Yang, X. Qi, et al., Fertility preservation in breast cancer patients: A systematic review and summary of current evidence and recommendations, *Breast Cancer Res. Treat.*, **215** (2026), 68. https://doi.org/10.1007/s10549-026-07901-z
5. A Hawazie, M Druce, Breast cancer risk and management in the endocrine clinic: A comprehensive review, *Clin. Endocrinol.*, 2026. https://doi.org/10.1111/cen.15209
6. A. Elengoe, A short review on breast cancer, *Int. J. Biotechnol. Biomed.*, **1** (2024), 1–11. https://doi.org/10.31674/ijbb.2024.v01i01.001
7. A. Husain, H. Khan, J. E. Fatima, S. A. Siddiqui, A. D. Singh, A. Husain, et al., Recent advancements in breast cancer therapies and biomarkers: Mechanisms and clinical significance, *Curr. Pharm. Biotechnol.*, **26** (2025), 2381–2402. https://doi.org/10.2174/0113892010343901250204054241
8. V. Velmurugan, B. Shanthakumar, M. K. Kathiravan, T. Sundarrajan, D. Priya, G. V. Anjana, Breast cancer epidemiology, risk factors, pathogenesis and advances in diagnosis and treatment: A review, In: *An Overview of Disease and Health Research*, **11** (2026), 54–81. https://doi.org/10.9734/bpi/aodhr/v11/7335

9. J. Kim, A. Harper, V. McCormack, H. Sung, N. Houssami, E. Morgan, et al., Global patterns and trends in breast cancer incidence and mortality across 185 countries, *Nat. Med.*, **31** (2025), 1154–1162. https://doi.org/10.1038/s41591-025-03502-3

10. *World Health Organization, Breast cancer*, World Health Organization, 2026. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

11. Z. Tao, A. Shi, C. Lu, T. Song, Z. Zhang, J. Zhao, Breast cancer: Epidemiology and etiology, *Cell Biochem. Biophys.*, **72** (2015), 333–338. https://doi.org/10.1007/s12013-014-0459-6

12. A. Din, Y. Li, Optimizing HIV/AIDS dynamics: Stochastic control strategies with education and treatment, *Eur. Phys. J. Plus*, **139** (2024), 812. https://doi.org/10.1140/epjp/s13360-024-05605-1

13. S. Kumar, R. Kumar, S. Momani, S. Hadid, A study on fractional COVID-19 disease model by using Hermite wavelets, *Math. Methods Appl. Sci.*, **46** (2023), 7671–7687. https://doi.org/10.1002/mma.7065

14. A. Din, Bifurcation analysis of a delayed stochastic HBV epidemic model: Cell-to-cell transmission, *Chaos Solitons Fract.*, **181** (2024), 114714. https://doi.org/10.1016/j.chaos.2024.114714

15. B. Ghanbari, S. Kumar, A study on fractional predator-prey-pathogen model with Mittag–Leffler kernel-based operators, *Numer. Meth. Part. D. E.*, **40** (2024), e22689. https://doi.org/10.1002/num.22689

16. P. J. Cook, R. Doll, S. A. Fellingham, A mathematical model for the age distribution of cancer in man, *Int. J. Cancer*, **4** (1969), 93–112. https://doi.org/10.1002/ijc.2910040113

17. A. M. Jarrett, E. A. Lima, D. A. Hormuth, M. T. McKenna, X. Feng, D. A. Ekrut, et al., Mathematical models of tumor cell proliferation: A review of the literature, *Expert Rev. Anticancer Ther.*, **18** (2018), 1271–1286. https://doi.org/10.1080/14737140.2018.1527689

18. P. M. Altrock, L. L. Liu, F. Michor, The mathematics of cancer: Integrating quantitative models, *Nat. Rev. Cancer*, **15** (2015), 730–745. https://doi.org/10.1038/nrc4029

19. R. P. Araujo, D. S. McElwain, A history of the study of solid tumour growth: The contribution of mathematical modelling, *Bull. Math. Biol.*, **66** (2004), 1039–1091. https://doi.org/10.1016/j.bulm.2003.11.002

20. D. S. Dixit, D. Kumar, S. Kumar, R. Johri, A mathematical model of chemotherapy for tumor treatment, *Adv. Appl. Math. Biosci.*, **3** (2012), 1–10.

21. G. Jordao, J. N. Tavares, Mathematical models in cancer therapy, *Biosystems*, **162** (2017), 12–23. https://doi.org/10.1016/j.biosystems.2017.08.007

22. G. Mahlbacher, L. T. Curtis, J. Lowengrub, H. B. Frieboes, Mathematical modeling of tumor-associated macrophage interactions with the cancer microenvironment, *J. Immunother. Cancer*, **6** (2018), 10. https://doi.org/10.1186/s40425-017-0313-7

23. N. Beerenwinkel, R. F. Schwarz, M. Gerstung, F. Markowetz, Cancer evolution: Mathematical models and computational inference, *Syst. Biol.*, **64** (2015), e1–e25. https://doi.org/10.1093/sysbio/syu081

24. M. Farhan, Z. Ling, J. Ding, Z. Shah, R. D. Dobrotă, A novel fractional computational neural framework for analyzing cancer model under chemotherapy drug, *Comput. Methods Biomech. Biomed. Eng.*, **29** (2026), 3192–3210. https://doi.org/10.1080/10255842.2025.2508227

25. A. Yin, D. J. A. Moes, J. G. van Hasselt, J. J. Swen, H. J. Guchelaar, A review of mathematical models for tumor dynamics and treatment resistance evolution of solid tumors, *CPT Pharmacometrics Syst. Pharmacol.*, **8** (2019), 720–737. https://doi.org/10.1002/psp4.12450

26. V. Gopakumar, A review on surrogate modelling and uncertainty quantification of neural differential equations, PhD Upgrade Viva Report, *Centre for Artificial Intelligence, Department of Computer Science University College London*, 2025.

27. M. A. Ullah, S. U. Saqib, Y. T. Shih, N. Abdullah, N. Becheikh, L. Kolsi, et al., A data-driven mathematical modeling framework for predicting HIV/HBV coinfection dynamics using machine learning approaches, *AIMS Math.*, **11** (2026), 22688–22723. http://doi.org/10.3934/math.2026916

28. S. U. Saqib, S. H. Fang, M. A. Z. Raja, K. S. Nisar, M. Shoaib, Design of intelligent neuro-supervised deep learning networks to analyze brain electrical activity rhythms of Parkinson's disease model, *Cogn. Neurodyn.*, **20** (2026), 32. https://doi.org/10.1007/s11571-025-10404-0

29. L. Monbroussou, M. Periyasamy, V. Kuzmin, P. Sekatski, V. Patapovich, A. Sagingalieva, et al., Hybrid quantum neural networks: Theory, implementations, and applications, *arXiv preprint arXiv:2608.01194*, 2026. https://doi.org/10.48550/arXiv.2608.01194

30. S. Kumar, S. Kumar, S. Momani, A deep neural network with optimization approach for a nonlinear fractional bird flu model with vaccination and treatment via spectral method, *Knowl.-Based Syst.*, **336** (2026), 115259. https://doi.org/10.1016/j.knosys.2026.115259

31. A. Din, Neural network assisted stability exploration of a stochastic Hepatitis B model incorporating dual transmission routes and immune response, *J. Appl. Math. Comput.*, **72** (2026), 116. https://doi.org/10.1007/s12190-026-02770-7

32. C. Banerjee, S. Kumar, S. Momani, A neural network optimization scheme for fractional economic and environmental mathematical model, *Int. J. Comput. Math.*, 2026, 1–39. https://doi.org/10.1080/00207160.2026.2637677

33. A. Din, Stochastic analysis and ANN-based approximation of measles transmission dynamics under Lévy noise, *AIMS Math.*, **11** (2026), 18715–18745. https://doi.org/10.3934/math.2026761

34. Y. Xu, Z. Shao, X. Liu, Y. Li, Z. An, C. Fu, et al., AI for science: Progress, challenges, and perspectives, *Innovation*, **7** (2026), 101530. https://doi.org/10.1016/j.xinn.2026.101530

35. S. U. Saqib, S. H. Fang, Y. T. Shih, N. Abdullah, N. Becheikh, L. Kolsi, et al., A novel hybrid neuro-stochastic framework for human immunodeficiency virus and Hepatitis B virus co-infection dynamics with vaccination: Computational aspects and dynamical analysis, with artificial neural network-based solutions, *Eng. Appl. Artif. Intell.*, **183** (2026), 116110. https://doi.org/10.1016/j.engappai.2026.116110

36. A. Muhamad, Quantum neural network in architectures, learning mechanisms, and emerging applications across domains: A review, *J. Multiscale Mater. Inform.*, **2** (2025), 30–39. https://doi.org/10.62411/jimat.v2i2.14929

37. A. K. Srivastava, S. Hussain, S. Sharma, A. Verma, N. Sharma, Quantum kernel-driven hybrid neural network for brain tumor classification, *IEEE Access*, **14** (2026), 47340–47354. https://doi.org/10.1109/ACCESS.2026.3677204

38. C. S. Chen, E. J. Kuo, Quantum adaptive self-attention for quantum transformer models, *Quantum Sci. Technol.*, **11** (2026), 035051. https://doi.org/10.1088/2058-9565/ae8b21

39. E. Alzahrani, M. M. El-Dessoky, M. A. Khan, Mathematical model to understand the dynamics of cancer, prevention diagnosis and therapy, *Mathematics*, **11** (2023), 1975. https://doi.org/10.3390/math11091975

40. X. Mao, *Stochastic Differential Equations and Applications*, 2 Eds, Horwood: Horwood Publishing, 2007.

41. F. Yu, M. Ullah, S. Ullah, M. Asiri, S. U. Khan, M. Qeshta, A multi-layer computational neural network approach for solving stochastic within-host Chikungunya model with an adaptive immune response, *Int. J. Biomath.*, 2025, 2550047. https://doi.org/10.1142/S1793524525500470

42. S. Li, M. Ullah, S. Ullah, T. Muhammad, Q. Iqbal, A novel stochastic neural network framework for modeling and simulation of within-host Chikungunya virus transmission with latency, *Knowl.-Based Syst.*, **329** (2025), 114412. https://doi.org/10.1016/j.knosys.2025.114412

43. J. Zhu, M. Ullah, S. Ullah, M. B. Riaz, A. B. Saqib, A. M. Alamri, et al., A novel numerical solution of nonlinear stochastic model for the propagation of malicious codes in wireless sensor networks using a high order spectral collocation technique, *Sci. Rep.*, **15** (2025), 228. https://doi.org/10.1038/s41598-024-82033-2



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)