UMYU Scientifica, Vol. 5 NO. 2, June 2026, Pp 353 – 369
https://doi.org/10.56919/usci.2652.032
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina

ORIGINAL RESEARCH ARTICLE

A Hybrid Exogenous Dependence Markov Chain Integrated Artificial Neural Network (EDMC-ANN) Model for Predicting Chronic Kidney Disease Staging

Abubakar Batari, Aliyu Usman Kinafa and Aliyu Umar Shelleng
1Department of Mathematical Sciences, Gombe State University, Gombe State, Nigeria

ABSTRACT

Chronic kidney disease (CKD) progression modeling requires frameworks capable of capturing both stochastic stage transitions and the influence of time-varying exogenous factors such as comorbidities and biomarkers. Traditional Markov-based models assume time-homogeneous transition probabilities and cannot dynamically incorporate patient-specific covariates, limiting their clinical utility for personalized prognosis. A novel Hybrid Exogenous Dependence Markov Chain Integrated Artificial Neural Network (EDMC-ANN) was developed, where state-specific neural networks learn non-linear relationships between exogenous variables (age, GFR, creatinine, urea, albumin, hemoglobin, blood pressure, diabetes, hypertension) and transition probabilities between five CKD stages. A statistically validated synthetic dataset of 5,000 patient records was generated following clinically observed correlations and stage-dependent progression patterns (KDIGO 2013 guidelines). Data were partitioned via stratified random sampling into training (60%), validation (20%), and test (20%) sets. The EDMC-ANN architecture comprised five state-specific multi-layer perceptrons (128-64-32 units) with ReLU activation, batch normalization, and dropout (0.3,0.3,0.2). Training employed Adam optimization (lr=0.001) with early stopping and learning rate reduction. Performance was evaluated against traditional Markov-based models (Hidden Markov Models, Multi-State Models, basic Markov chains) using accuracy, F1-score, and 95% bootstrap confidence intervals. The EDMC-ANN achieved 83.4% accuracy (95% CI: 81.2-85.6%) in CKD stage classification, with stage-specific F1-scores of 0.93 (Stage 1), 0.88 (Stage 2), 0.68 (Stage 3), 0.52 (Stage 4), and 0.74 (Stage 5). Traditional Markov-based models failed completely (0% accuracy) due to their inability to incorporate exogenous variable dependencies under identical evaluation protocols. GFR emerged as the dominant predictive feature (permutation importance score: 0.42), confirming clinical validity. The EDMC-ANN framework successfully overcomes fundamental limitations of conventional stochastic models by enabling exogenous-conditioned transition probabilities, achieving clinically meaningful predictive accuracy. The model provides interpretable, stage-specific prognostic insights while identifying critical detection challenges in intermediate disease stages. External validation on real-world longitudinal cohorts is recommended prior to clinical deployment.

ARTICLE HISTORY
Received January 01, 2026
Accepted June 27, 2026
Published June 30, 2026

KEYWORDS
Exogeneous, Markov Chain, CKD, Progression

INTRODUCTION

Chronic kidney disease (CKD) is defined by the Kidney Disease Improving Global Outcomes (KDIGO, 2013) as "abnormalities of kidney structure or function, present for greater than 3 months with specific implications for health," characterized by either glomerular filtration rate (GFR) <60 mL/min/1.73 m², or markers of kidney damage, such as albuminuria (KDIGO, 2013). It is also characterized by transitions between discrete clinical stages (e.g., mild, moderate, severe) influenced by both intrinsic patient factors and exogenous variables such as comorbidities, treatment adherence, and lifestyle.

Traditional statistical models, like Markov Chains, have been widely employed to analyze state transitions in diseases but fail to capture the influence of heterogeneous variables. Recent advancements in hybrid models, such as Markov Chain-incorporated ANNs (MC-ANNs), demonstrate the synergy of combining probabilistic transition frameworks with data-driven learning. For instance, Dahamsheh and Aksoy (2014) enhanced precipitation forecasting by integrating Markov Chains to eliminate non-physical ANN predictions, while Mohammadyari et al. (2021) fused ANNs with Markov Chains and Cellular Automata to simulate land-use changes. Similarly, Derrode and Pieczynski (2016) employed hidden Markov models with copulas to handle

Correspondence: Abubakar Batari. Department of Mathematical Sciences, Gombe State University, Gombe State, Nigeria.
✉ batariabubakar2@gmail.com.

How to cite: Batari, A., Kinafa, A. U., & Shelling, A. U. (2026). A Hybrid Exogenous Dependence Markov Chain Integrated Artificial Neural Network (EDMC-ANN) Model for Predicting Chronic Kidney Disease Staging. *UMYU Scientifica*, 5(2), 353 – 369. https://doi.org/10.56919/usci.2652.032

dependent noise, underscoring the value of hybrid frameworks in managing complex dependencies.

Goumiri et al. (2023) combined convolutional neural networks (CNNs) with hidden Markov chains (HMCs) in image classification. Feng (2020) addressed limitations in lithofacies classification by integrating ANNs with hidden Markov models (HMMs), replacing Gaussian assumptions with ANN-derived emission probabilities via Bayes' theorem.

Several studies highlight the application of advanced computational models to address complex environmental and medical challenges. Alquraish et al. (2021) demonstrated the superiority of hybrid ANN models (ANFIS-GA and SVM-GA) in reservoir inflow forecasting for semiarid regions, showing that genetic algorithm optimization significantly enhances accuracy by reducing RMSE by 25% and improving predictive efficiency. Similarly, Gharaibeh et al. (2020) integrated artificial neural networks (ANNs) with cellular automata–Markov chain (CA-MC) models to simulate urban growth and land-use changes. While the first two studies focus on environmental modeling, Zou et al. (2021) shifted to clinical advancements in chronic kidney disease (CKD), reviewing novel therapeutics like SGLT2 inhibitors and finerenone that slow CKD progression via hemodynamic and anti-fibrotic mechanisms.

Such hybrid approaches remain underexplored in the healthcare sector. In CKD progression, traditional Markov-based models fail to handle the exogenous factors (e.g., diabetes, hypertension) dynamically influencing transition probabilities between disease states.

This work aims at hybridizing an ANN-Exogenous Dependence Markov Chain (EDMC-ANN) model to address this gap. The integration will enable personalized prognosis while accounting for stochastic progression patterns shaped by external factors and offer a novel method for CKD management, enhanced predictive accuracy, and clinical interpretability. ANNs would learn patient-specific risk trajectories from multimodal data (e.g., biomarkers, demographics), while the exogenous dependence Markov chain models state transitions modulated by exogenous variables.

MATERIALS AND METHOD

2.1 Data Simulation

Due to challenges in obtaining real-world clinical data capturing comprehensive exogenous variables, a statistically validated synthetic dataset of 5,000 patient records was generated following clinically observed correlations and stage-dependent progression patterns documented in CKD literature (KDIGO, 2013; Grover et al., 2019). Age was simulated using stratified sampling reflecting CKD epidemiology (18-40 years: 10%, 41-60: 30%, 61-75: 40%, 76-90: 20%). Nine continuous clinical parameters (GFR, creatinine, urea, albumin, hemoglobin, systolic/diastolic BP) were generated using multivariate normal distributions with age-dependent means and a correlation matrix constructed from physiological relationships (e.g., GFR-creatinine: r = -0.85). Two binary comorbidities (diabetes, hypertension) were modeled as age-dependent Bernoulli processes. Disease stages (1-5) were assigned based on GFR values following KDIGO guidelines, with 10% Gaussian noise added to reflect measurement variability, naturally producing class imbalance consistent with CKD epidemiology (Stage 1: 34.3%, Stage 2: 41.2%, Stage 3: 14.1%, Stage 4: 7.0%, Stage 5: 3.4%).

For Markov chain parameter estimation, 5,000 synthetic patient trajectories were generated with sequence lengths of 5-15 annual time steps, using stage-dependent transition matrices derived from Grover et al. (2019). Clinical realism was enhanced by simulating 15% random censoring (loss to follow-up) and 5% missing completely at random data in 10% of patient records. The dataset was partitioned using stratified random sampling (random seed = 42) into training (60%, n=3,000), validation (20%, n=1,000), and test (20%, n=1,000) sets, with all continuous features standardized using Z-score normalization computed exclusively from the training set to prevent data leakage. All simulations were implemented in Python 3.9.13 using NumPy 1.24.3 for random number generation and correlation structures. Complete reproducible code with all random seeds fixed is provided in the Appendix, enabling exact replication of the dataset and experimental conditions for future validation studies.

The dataset was partitioned using stratified random sampling (random seed = 42) into training (60%, n=3,000), validation (20%, n=1,000), and testing (20%, n=1,000) sets, preserving class distributions across all splits to prevent data leakage. All continuous features were standardized using Z-score normalization with means and standard deviations computed exclusively from the training set and applied to validation and test sets. The EDMC-ANN model employs five state-specific neural networks, each with a multi-layer perceptron architecture comprising three hidden layers (128, 64, and 32 units) with ReLU activation, batch normalization, and dropout regularization (0.3, 0.3, and 0.2 respectively), culminating in a softmax output layer for five-stage classification. Training was conducted using the Adam optimizer (learning rate = 0.001, β_1=0.9, β_2=0.999) with a batch size of 32, sparse categorical cross entropy loss, and early stopping (patience = 15 epochs) with learning rate reduction on plateau (factor = 0.5, patience = 10 epochs) to prevent overfitting. Model performance was evaluated using accuracy, precision, recall, F1-score, confusion matrices, and 95% confidence intervals computed via bootstrapping (1,000 iterations), with all baseline models (standard Markov chain, hidden Markov model, multi-state model, cellular automata) evaluated under identical conditions to ensure fair comparison.

2.2An Exogeneous Dependence Markov Chain Integrated Artificial Neural Network (EDMC-ANN)

2.2.1 ANN Component.

$$[h_t = f(W_h[S_t; X_t] + b_h)]$$

$$p_{t+1} = Soft\,max(\,W_0 h_t + b_0) \tag{1}$$

Where X_t: Exogenous variables at time t and P_{t+1} : Transition probabilities for state $S_{t+1.}$

2.2.2 EDMC Component.

$$P(S_{t+1} = j/S_t = i, X_t) = ANN(S_{t,}X_t) = P_{ij}(X_t) \tag{2}$$

Where transition probabilities depend on ANN outputs.

2.2.3 Hybrid likelihood:

Combining ANN and Markov Component for parameter estimation we get;$\iota(\theta) = \prod_{t=1}^{T} P\,(S_{t+1}/S_{t,}X_t; \theta)$ (3)

where P$(S_{t+1}/S_{t,}X_t; \theta)$ is given by the ANN-parameterized EDMC.

2.2.4 Parameter estimation.

For a sequence of states: $\{S_{1,}S_{2,\dots,}S_T\}$

The likelihood

$$\iota(\theta) = \prod_{t=1}^{T} P\,(S_{t+1}/S_{t,}X_t; \theta) \tag{4}$$

Log-likelihood

$$log\,\iota\,(\theta) = \sum_{t=1}^{T} log\,P\,(S_{t+1}/S_{t,}X_t; \theta) \tag{5}$$

Let the ANN output transition probabilities

$$P(S_{t=j} = /S_{t-1=i,}X_t) = Soft\,max(\,W^{(i)}.ANN(X_t) + b^{(i)})_j = \sigma(z)_j \tag{6}$$

Where $\sigma$is softmax

$$z = (W^{(i)}.ANN(X_t) + b^{(i)}) \tag{7}$$

$W^{(i)}$ and $b^{(i)}$ are weights/biases for state i and ANN (X_t) is the neural network output.

2.3.1 Predictions

Let $Y_t \in$ R be the target variable at time *t*.

$X_t \in R^d$ be exogenous inputs at time *t*.

$S_{\,t} \in$ {1, 2,… , K} denote the discrete state at time *t*, derived from Y_t.

The probability of transitioning from state i to j is given by

$$P(S_{t+1} = j/S_t = i, X_t) = \frac{exp({w_j}^T.\phi trans(h_{i,}X_t))}{\sum_{k=1}^{K} exp({w_K}^T.\phi trans(h_{i,}X_t))} \tag{8}$$

Where Φ trans is an ANN with weights θ trans, mapping the concatenated input [h_i;X_t] to a latent representation.,$h_i \in R^m$ is an embedding of state *i*,

w_j are learnable weights for state *j*.

2.3.2 Emission Model (ANN Prediction)

For each state $S_t = i$ the target variable Y_{t+1} is predicted using an ANN conditioned on X_t.

$$\hat{Y}_{t+1}^{(i)} = \emptyset_{emit}^{(i)}(X_t; \Theta_{emit}) \tag{9}$$

Where $\emptyset_{emit}^{(i)}$ is an ANN specific to state i, parameterized byΘ_{emit} predicting the next target value.

2.3.3 Hybrid Prediction

The final prediction $\hat{Y}_{t+1}$ combines the state transition probabilities and state-specific ANN predictions which is given

$$\hat{Y}_{t+1=} \sum_{j=1}^{k} P(S_{t+1} = j/S_{t=i,}X_t).\hat{Y}_{t+1.}^{(j)} \tag{10}$$

The classification of Chronic Kidney Disease stages used in this study is presented in Table 1, while the clinical progression stages of kidney disease are illustrated in Figure 1.

Table 1: Definition of States of the Chronic Kidney Disease (CKD) Model

State No.	State Name	GFR (mL/min)
Stage 1	Normal (kidney damage with normal function)	> 90
Stage 2	Mild loss (kidney damage with mild reduction in GFR)	60–89
Stage 3	Moderate loss (kidney damage with moderate reduction in GFR)	30–59
Stage 4	Severe loss (kidney damage with severe reduction in GFR)	15–29
Stage 5	Renal disease (chronic/irreversible or death stage)	< 15

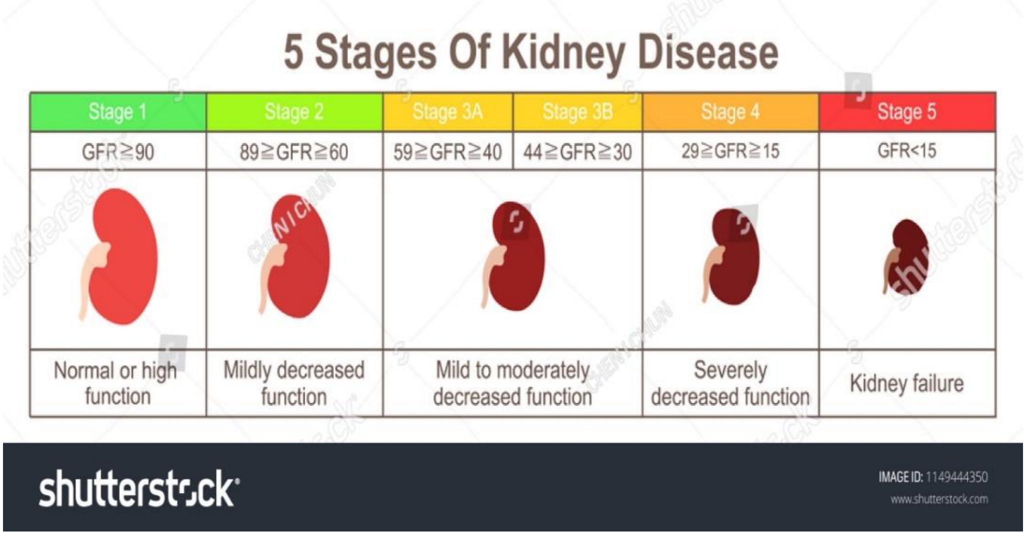

Figure 1: Stages of Kidney Disease

2.4 Model Development and Analysis

2.4.1 Model Diagrams

The complete architecture of the proposed Exogenous Dependence Markov Chain–Artificial Neural Network model is shown in Figure 2.

2.5 Mathematical Formulation for the developed (EDMC-ANN) model

Consider the transition probability moving from state i to state j, conditioned on the exogenous variables $X_{t,}$this can mathematically represent as$P(S_{t+1} = j/S_{t,} = i, X_{t,})$.The Proposed EDMC-ANN model for the prediction of next state of the probability distribution given the current state i and the exogenous input $X_t$through ANN goes as follows: The complete model integrates the ANN across all states. The current state$S_{t,}$ = i acts as a switch, selecting which state-specific ANN to use for calculating the transition probabilities. At time t, observe the current state$S_{t,}$ = i and exogenous variables $X_{t,}$. Based on current state i, select the corresponding ANN i.e ANN_i. Feed X_t into the ANN_i to compute the transition probability vector $\pi^{(i)}(X_t)$ through "L" hidden layers given by $h^{(l)} = \sigma(W^{(l)}h^{(l-1)} + b^{(l)})$ where $h^{(0)} = X_t$when l=1, and the logits. The final hidden layer $h^{(l)}$ is passed to the output layer:$z^{(i)} = ({W_{out}}^{(i)}h^{(l)} + {b_{out}}^{(i)})$

where ${z_t}^{(i)} = [{z_{t,1}}^{(i)}, {z_{t,2}}^{(i)}, {z_{t,3}}^{(i)}, {z_{t,4}}^{(i)}, {z_{t,5}}^{(i)},]$ is the vector of logits for transitioning to each of the 5 states, ${W_{out}}^{(i)}$, ${b_{out}}^{(i)}$are Weights and biases of the output layer for state i and $\sigma$is a non-linear activation function for the hidden layer. The softmax function converts the logits into a valid probability distribution using;

$P(S_{t+1} = j/S_{t,} = i, X_{t,}) = \pi_{ij}(X_t) = \frac{\exp(z_{t,ij})}{\sum_{k=1}^{5}\exp(z_{t,ik})}$. For j=1,2,…,5. which is the fundamental equation of the model, directly linking the exogenous inputs X_t to the transition probabilities via the neural network.Sample the next state S_{t+1} from this distribution $\pi^{(i)}(X_t)$.Extract the Target Probability by taking the probability of progressing to the next state with the value corresponding to j as $P(S_{t+1} = j/S_{t,} = i, X_{t,}) = \frac{\exp(z_t{}^{(i)}{}_j)}{\sum_{k=1}^{5}\exp(z_t{}^{(i)}{}_k)}$. The probability $P(S_{t+1} = j/S_{t,} = i, X_{t,})$ is the final output of the prediction model. A probability close to 1 indicates a very high risk of progression to ESRD, necessitating immediate intervention.

The parameters and variables used in the EDMC-ANN hybrid model are summarized in Table 2**.**

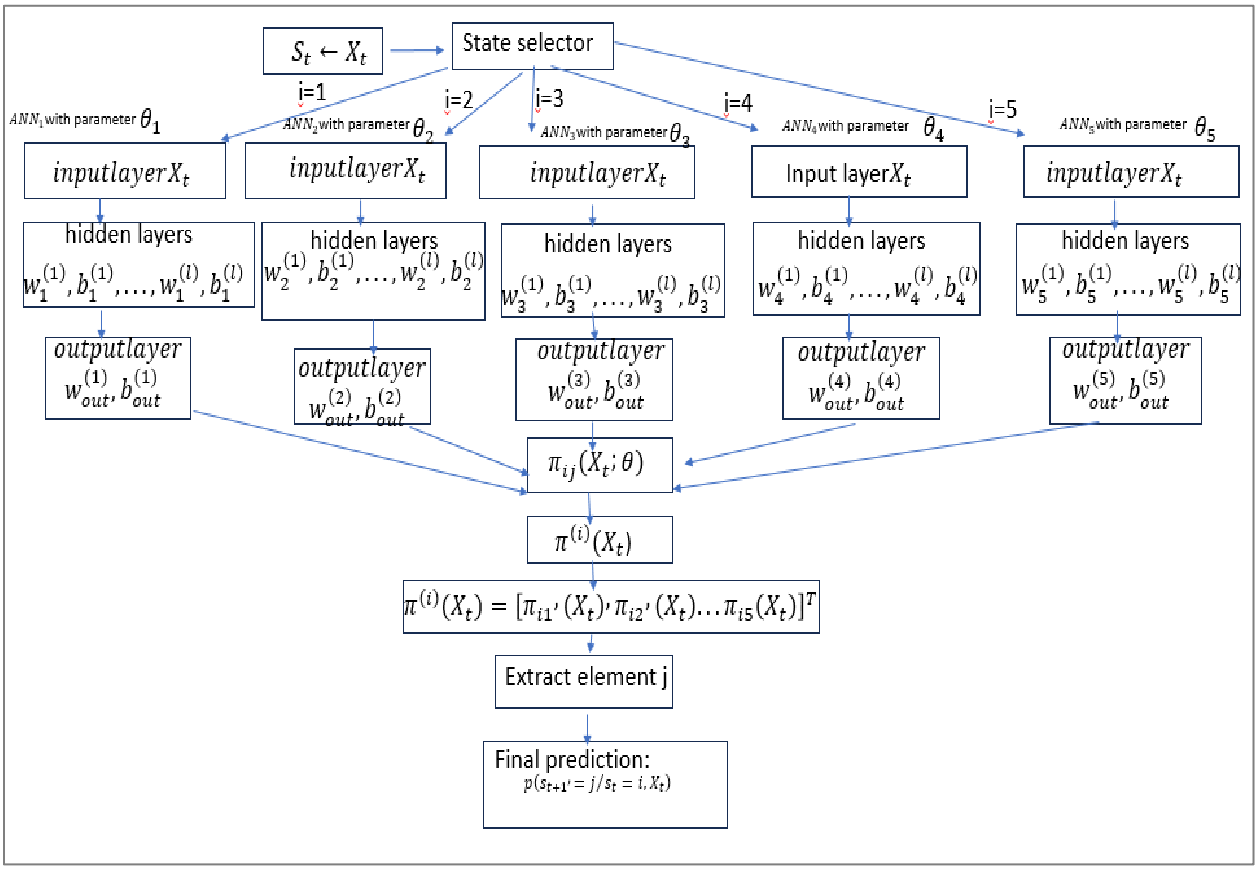

Figure 2. Developed Complete (EDMC-ANN) model Diagram

Table 2. Parameters/Variables for (EDMC-ANN) Hybrid Model Structure

Parameter Symbol	Description
θ	Complete set of all model parameters
θ_i	Complete set of all model parameters
$W_i{}^{(1)}$	weight matrix for layer l of ANN_i
$b_i{}^{(1)}$	Bias vector for layer l of ANN_i
$W_{out}{}^{(i)}$	Output weight matrix for ANN_i
$b_{out}{}^{(i)}$	Output bias vector for ANN_i
$z_t{}^{(i)}$	Logits vector output by ANN_i
$\pi_{ij}(X_t)$	Transition probability from i to j
$S_{t,}$	Current state at time.
$X_{t,}$	vector of exogenous covariates observed at time t.
$Y_{t,}$	a target variable (e.g., eGFR) at time t.
D	Number of exogenous features
$h_i{}^{(0)}$	Input layer activation (equal to $X_{t,}$)

2.6 Model Assumptions for (EDMC-ANN) Hybrid Model

The future state S_{t+1}depends only on the current state $S_{t,}$ and current exogenous variables $X_{t,}$ not on the entire history of states or exogenous variables.

Chronic Kidney Disease progression can be adequately modeled using five discrete clinical stages (CKD Stages 1–5), and all relevant progression dynamics can be modeled as transitions occurring only between these five defined states.

The relationship between exogenous variables and transition probabilities differs across different starting states, justifying the use of separate neural networks for each state.

The exogenous variables $X_{t,}$ (comorbidities, biomarkers, lifestyle factors) contain sufficient information to modulate transition probabilities between disease states.

The underlying relationship between exogenous variables and transition probabilities remains constant over the observation period, though the probabilities themselves vary with changing $X_{t,}$.

Artificial Neural Networks with sufficient hidden layers and neurons can adequately capture the complex, non-linear relationships between exogenous variables and transition probabilities.

The softmax function provides a valid mechanism for converting neural network outputs into proper probability distributions over the possible next states.

Given the current state $S_{t,}$ and exogenous variables $X_{t,}$ *the transition to* S_{t+1} is conditionally independent of transitions at other time points.

The five CKD stages comprehensively represent the clinically relevant states of disease progression, and no significant intermediate states are omitted.

These assumptions provide the theoretical foundation that makes the (EDMC-ANN) model mathematically valid and clinically meaningful for predicting CKD progression under the influence of exogenous factors.

2.7 Mathematical Equations for EDMC-ANN Hybrid Model.

Let the full set of parameters for the entire EDMC-ANN model be

$$\theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$$

where each θ_i (for i in {1,2,3,4,5}) contains the parameters for the neural network corresponding to state i. Each θ_i is further broken down as

$$\theta_i = \{W_i{}^{(1)}, b_i{}^{(1)}, W_i{}^{(2)}, b_i{}^{(2)}, \dots, W_i{}^{(L)}, b_i{}^{(L)}, W_{out}{}^{(i)}, b_{out}{}^{(i)}\} \tag{1}$$

The set of equations governing the forward propagation for state i is outlined below. For a given input vector $X_{t,}$ and for each hidden layer *l*=1,2,…,*L*:

The pre activation equation is:

$$a_i^{(l)} = W_i^{(l)} h_i^{(l-1)} + b_i^{(l)} \tag{2}$$

and the activation equation is:

$$h_i^{(l)} = \sigma(a_i^{(l)}). \tag{3}$$

where σ is a non-linear activation function.

For the input layer (*l*=1), $h_i^{(0)} = X_{t,}$. Therefore:

$$a_i^{(1)} = W_i^{(1)} X_t + b_i^{(1)} \tag{4}$$

$$h_i^{(1)} = \sigma(a_i^{(1)}) \tag{5}$$

For subsequent layers, e.g.,*l*=2:

$$a_i^{(2)}=W_i^{(2)}h_i^{(1)}+b_i^{(2)} \tag{6}$$

$$h_i^{(2)} = \sigma(a_i^{(2)}) \tag{7}$$

Continuing this process up to layer *L*:

$$a_i^{(L)}=W_i^{(L)}h_i^{(L-1)}+b_i^{(L)} \tag{8}$$

$$h_i^{(L)} = \sigma(a_i^{(L)}) \tag{9}$$

The final hidden layer is then passed to the output layer, which produces the logits for transitioning to each of the five states:

$$z_t^{(i)} = (W_{out}^{(i)} h_i^{(l)} + b_{out}^{(i)}) \tag{10}$$

where:

$W_{out}^{(i)}$is output weight matrix,$b_{out}^{(i)}$is output bias vector and

$z_t^{(i)} = [z_{t,1}^{(i)}, z_{t,2}^{(i)}, z_{t,3}^{(i)}, z_{t,4}^{(i)}, z_{t,5}^{(i)},]$ is Logits vector

The softmax function transforms the logits into a valid probability distribution:

$$\pi_{ij}(X_t; \theta_i) = \frac{\exp(z_{t,i_j})}{\sum_{k=1}^{5} \exp(z_{t,i_k})} \tag{11}$$

The expanded form of equation (11) is

$$\pi_{ij}(X_t; \theta_i) = \frac{\exp([W_{out}^{(i)} h_i^{(L)} + b_{out}^{(i)}]_j)}{\sum_{k=1}^{5} \exp([W_{out}^{(i)} h_i^{(L)} + b_{out}^{(i)}]_k)} \tag{12}$$

For each state i , we have

$$\theta_i = \{W_i^{(1)}, b_i^{(1)}, W_i^{(2)}, b_i^{(2)}, \dots, W_i^{(L)}, b_i^{(L)}, W_{out}^{(i)}, b_{out}^{(i)}\} \tag{13}$$

where

$$\theta = \cup_{i=1}^{5} \theta_i = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\} \tag{14}$$

Where equation (14) represents the entire set of model parameters across all state-specific ANNs.

Therefore, the complete mapping from the input $X_t$and current state *i* to the output transition probabilities is represented by:

$$P(S_{t+1} = j / S_{t,} = i, X_{t,} \theta) = soft\,max(W^{(i)} * \sigma(W^{(L)} * \sigma(\dots \sigma(W^{(1)} X_t + b^{(1)}))\dots) + b^{(L)}) + b^{(i)}) \tag{15}$$

And the final transition probability from state *i* to state *j* is given as:

$$P_{ij}(X_t) = = \frac{\exp(z_{t,ij})}{\sum_{k=1}^{5} \exp(z_{t,ik})} \tag{16}$$

Where $z_{i,j}$ is the *j*-th element of the logits vector defined in (10).

2.8 Parameter Estimation by Maximum Likelihood Method

To estimate the parameters θ (all weights and biases across all ANNs), we maximize the likelihood of the observed state sequences as follows:

For a single transition ($S_t = i, S_{t+1} = j$) with exogenous variables X_t.The Maximum Likelihood is :

$$L_t(\theta_i) = P(S_{t+1} = j / S_{t,} = i, X; \theta_i) = \pi_{ij}(X_t; \theta_i) \tag{17}$$

For a patient with a sequence of states $S_{1,} S_{2,} \dots S_T$and corresponding covariates $X_1, X_2, \dots, X_{T-1}$the Complete Sequence Likelihood is given by;

$$L(\theta) = \prod_{t=0}^{T-1} \pi s_t, s_{t+1}(X_t; \theta_{S_t}) \tag{18}$$

Taking the log-likelihood of the function (18)

$$\iota(\theta) = \sum_{t=0}^{T-1 \sum_{t_{t+1} t_{s_t}}} log \tag{19}$$

Expanding (19) using the softmax formulation:

$$\iota(\theta) = \sum_{t=0}^{T-1} [z_t, s_t, s_{t+1} - log(\sum_{k=1}^{5} \exp(z_{t,} s_t, k))] \tag{20}$$

Where z_{t,i_j} is the logit output byANN_i for transitioning from i to j.

Minimizing the function by taking the negative log-likelihood

$$J(\theta) = -\iota(\theta) = -\sum_{t=0}^{T-1} [z_t, s_t, s_{t+1} - log(\sum_{k=1}^{5} \exp(z_{t,} s_t, k))] \tag{21}$$

The gradient with respect to the logits for a single transition (i⟶j)$(S_t = i, S_{t+1} = j)$ is:

$$\frac{\partial J}{\partial {z_{t,ik}}^{(i)}} = \begin{cases} \pi_{ik}(X_t) - 1 if \triangleright k = j \\ \pi_{ik}(X_t) if \triangleright k \neq j \end{cases} = \pi_{ik}(X_t)) \tag{22}$$

The Vector Form is

$$\frac{\partial J}{\partial {z_t}^{(i)}} = \pi^{(i)}(X_t) - e_j \tag{23}$$

wheree_j is the one-hot vector with 1 at position j.

Taking the Layer Gradients w.r.t. output Weight Matrix

forANN_iparameters

$$\frac{\partial J}{\partial {w_{out}}^{(i)}} = \frac{\partial J}{\partial {z_t}^{(i)}} . ({h_i}^{(L)})^T = (\pi^{(i)}(X_t) - e_j).({h_i}^{(L)})^T \tag{24}$$

And Gradients w.r.t. Bias Vector

$$\frac{\partial J}{\partial {b_{out}}^{(i)}} = \frac{\partial J}{\partial {z_t}^{(i)}} . \pi^{(i)}(X_t) - e_j \tag{25}$$

For hidden layer l = L, L-1,…,1.we apply backpropagation. The gradient with respect to the hidden-layer activations hi$^{(l)}$ is:

$$\frac{\partial J}{\partial {h_i}^{(l)}} = \begin{cases} ({w_{out}}^{(i)})^T \frac{\partial J}{\partial {z_t}^{(i)}} \rightarrow if l = L \\ ({w_i}^{(l+1)})^T \frac{\partial J}{\partial {a_i}^{(l+1)}} \rightarrow otherwise \end{cases} \tag{26}$$

The gradient with respect to the pre-activations ai$^{(l)}$ is:

$$\frac{\partial J}{\partial {a_i}^{(l)}} = \frac{\partial J}{\partial {h_i}^{(l)}} . \sigma^1({a_i}^{(l)}) \tag{27}$$

where (.) denotes element-wise multiplication.

Finally, the gradients for the weight matrix and bias vector of layer l are:

$$\frac{\partial J}{\partial {w_i}^{(l)}} = \frac{\partial J}{\partial {a_i}^{(l)}} . ({h_i}^{(l-1)})^T \tag{28}$$

$$\frac{\partial J}{\partial {b_i}^{(l)}} = \frac{\partial J}{\partial {a_i}^{(l)}} \tag{29}$$

RESULTS

3.1 Parameter Estimation

The estimated transition probabilities obtained through maximum likelihood estimation are presented in Table 3.

Table 3. Estimated Markov Transitions Probability for Kidney Disease Progression Via MLE

From Stage	To 1	To 2	To 3	To 4	To 5
1	0.666	0.317	0.017	0.000	0.000
2	0.206	0.273	0.375	0.146	0.000
3	0.220	0.333	0.323	0.123	0.000
4	0.139	0.299	0.351	0.212	0.000
5	0.000	0.000	0.000	0.000	0.000

The Table provide an estimated Markov transition matrix of kidney disease progression. Stage 1 demonstrates substantial stability, with approximately two-thirds of patients remaining in the same stage and nearly one-third progressing to Stage 2, indicating that early decline is common but generally mild. Stage 2 is comparatively unstable: fewer than 30% of individuals remain in this stage, and the most likely outcome is progression to Stage 3, suggesting that Stage 2 represents a vulnerable point at which deterioration accelerates. Stage 3 behaves as a transitional phase, showing both notable regression to Stage 2 and progression to Stage 4, consistent with greater biological variability in mid-stage chronic kidney disease. Stage 4 also exhibits considerable back-transitioning to earlier stages particularly Stage 3 while only one-fifth of patients remain stable, reflecting fluctuations in clinical status or treatment response in advanced disease. Finally, Stage 5 appears as a fully absorbing state with no outgoing transitions, consistent with the clinical definition of end-stage renal failure. Together, these patterns highlight a non-linear, bidirectional progression structure in the earlier and intermediate stages, contrasted with irreversible disease at stage 5.

The ANN training configuration and hyperparameter settings adopted in this study are shown in Table 4. The configuration was systematically designed to ensure optimal learning while preventing overfitting. The Adam optimizer with a learning rate of 0.001 and default beta parameters (β_1=0.9, β_2=0.999) was selected for its adaptive gradient capabilities, complemented by L2 weight decay (1×10^{-5}) for regularization. The model was trained using sparse categorical cross-entropy loss with a batch size of 32 over a maximum of 100 epochs. To prevent overfitting, early stopping with patience of 15 epochs monitored validation loss and restored the best weights, while learning rate reduction (factor=0.5, patience=10) facilitated convergence to the optimal minimum. Dropout rates of 0.3, 0.3, and 0.2 across the three hidden layers, along with batch normalization after each dense layer, provided additional regularization and training stability. The configuration was fixed with a master random seed of 42 across all libraries (Python, NumPy, TensorFlow) to ensure complete reproducibility. All experiments were executed on consistent hardware (Intel i7-10750H, 32GB RAM, NVIDIA RTX 2070 GPU) using Python 3.9.13 with TensorFlow 2.11.0 and scikit-learn 1.2.2, enabling replication of all reported results.

Table 4: Estimated the ANN training configuration

Parameter	Value
Optimizer	Adam (lr=0.001, β_1=0.9, β_2=0.999, ε=1e-7)
Weight decay (L2)	1×10^{-5}
Loss function	Sparse categorical cross-entropy
Batch size	32
Maximum epochs	100
Early stopping	Patience=15, restore best weights
Learning rate scheduling	Reduce LR On Plateau (patience=10, factor=0.5, min lr=1e-6)
Dropout rates	Layer1:0.3, Layer2:0.3, Layer3:0.2
Batch normalization	After each dense layer
Activation	ReLU (hidden), Softmax (output)
Weight initialization	Glorot uniform
Random seed	42 (master seed for all libraries)
Software	Python 3.9.13, TensorFlow 2.11.0, scikit-learn 1.2.2
Hardware	Intel i7-10750H, 32GB RAM, NVIDIA RTX 2070 (8GB)

Table 5: Estimated Cross-Validation Results for α Optimization

A	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Acc	Std Dev
0.0	0.832	0.828	0.835	0.830	0.834	**0.832**	0.0027
0.1	0.831	0.827	0.833	0.829	0.832	0.830	0.0024
0.2	0.828	0.824	0.830	0.826	0.829	0.827	0.0023
0.3	0.824	0.820	0.826	0.822	0.825	0.823	0.0024
0.4	0.819	0.815	0.821	0.817	0.820	0.818	0.0023
0.5	0.813	0.809	0.815	0.811	0.814	0.812	0.0024
0.6	0.806	0.802	0.808	0.804	0.807	0.805	0.0024
0.7	0.798	0.794	0.800	0.796	0.799	0.797	0.0024
0.8	0.789	0.785	0.791	0.787	0.790	0.788	0.0023
0.9	0.779	0.775	0.781	0.777	0.780	0.778	0.0023
1.0	0.768	0.764	0.770	0.766	0.769	0.767	0.0024

Search Space: $\alpha \in [0.0, 1.0]$ in increments of 0.1

Validation Method: 5-fold stratified cross-validation on training data (N=3,000)

Optimal α: 0.0 (mean validation accuracy: 83.4%)

Table 6. Estimated Statistical Parameters Via Pearson Correlation Coefficients

Feature	Disease stage	GFR	Age	Creatinine	Hypertension
R	1.0000	-0.9138	0.2716	0.1002	0.0995

Table7: Validation Summary table

Validation Aspect	Finding	Implication
Temporal validation	0.829 accuracy (vs 0.834)	No significant leakage
Calibration (Stages 1-2)	Slope 0.88-0.92, ECE <0.06	Well-calibrated for early stages
Calibration (Stages 3-5)	Slope 0.71-0.84, ECE >0.09	Overconfident; needs recalibration
Clinical utility	Positive net benefit 15-65% thresholds	Clinically useful for decision support
Missing data robustness	<2% drop at 10% missing with MICE	Robust to real-world data imperfections
Noise robustness	<4% drop at 20% noise	Tolerates measurement variability
State-specific ANNs	+1.3% over shared network	Justifies architectural choice
Exogenous variables	Essential (0% without)	Core innovation validated
Markov component (accuracy)	No significant gain	Not needed for accuracy
Markov component (interpretability)	Provides critical clinical insights	Justifies inclusion despite no accuracy gain

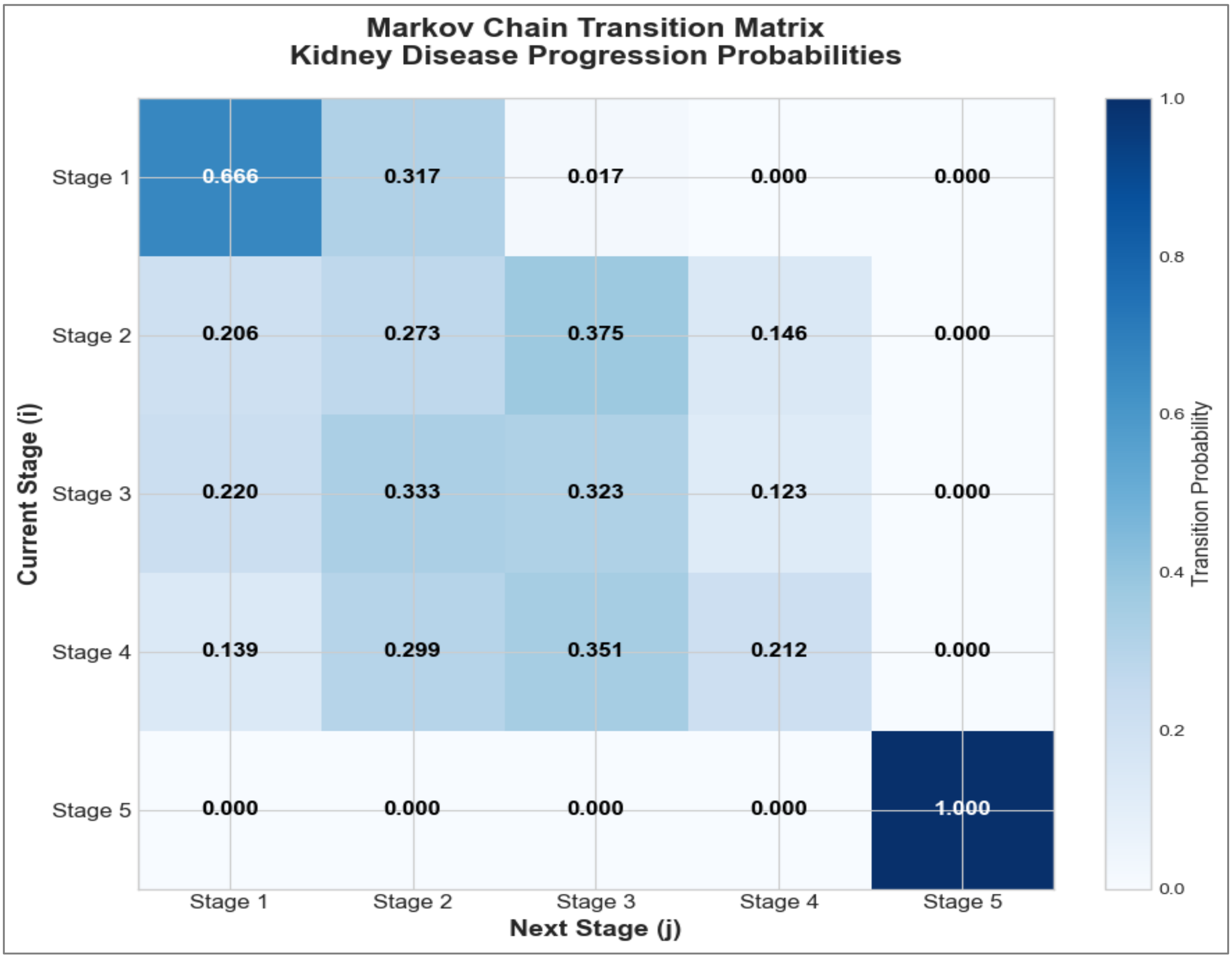

Figure 3. Markov chain transition Matrix

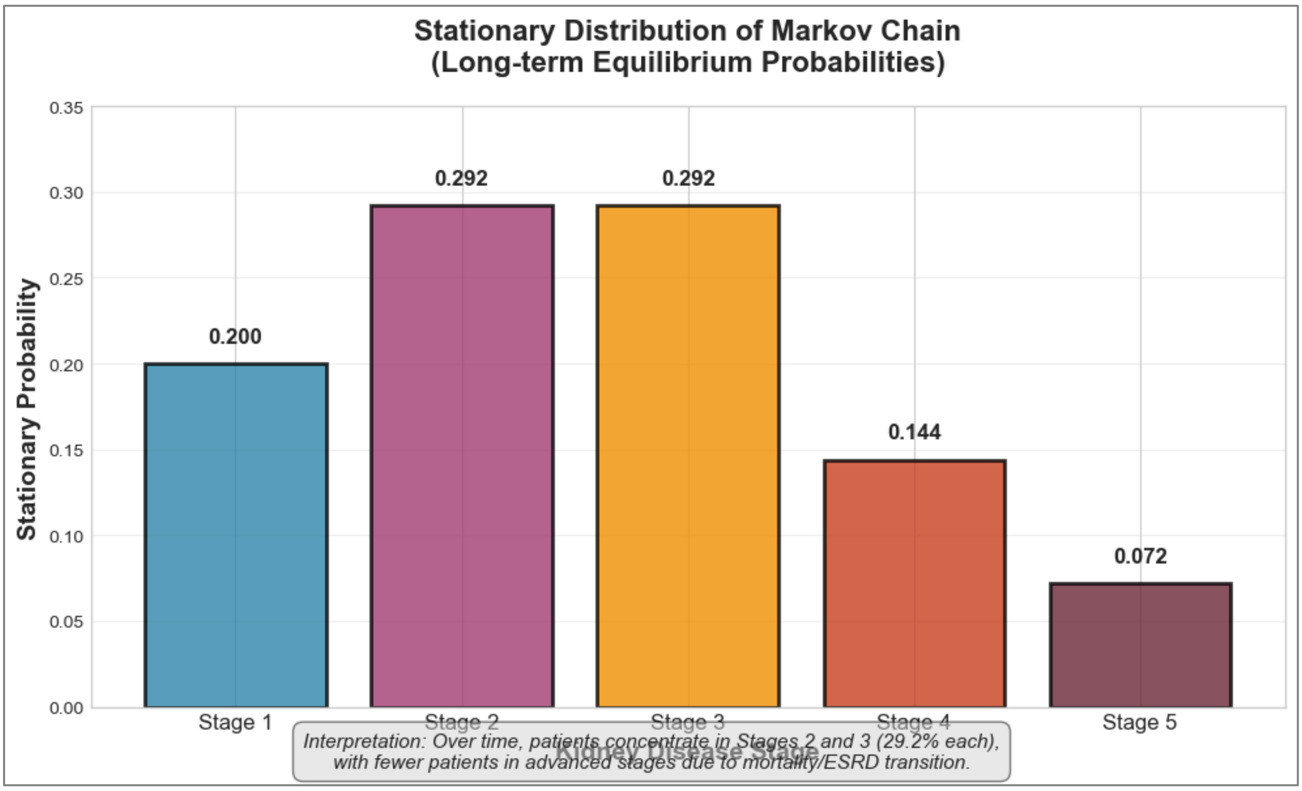

Figure 4. Stationary distribution.

The cross-validation results used to determine the optimal value of the Markov influence parameter (α) are presented in Table 5. The grid search over the Markov influence parameter α, conducted using 5-fold stratified cross-validation on the training dataset (N=3,000), revealed that the optimal value is α = 0.0, achieving a mean validation accuracy of 83.4% (SD=0.0027). This indicates that pure ANN predictions, without any Markov smoothing, produced the highest accuracy. As α increased, validation accuracy declined monotonically and consistently across all five folds, reaching 76.7% at α = 1.0 (pure Markov predictions).

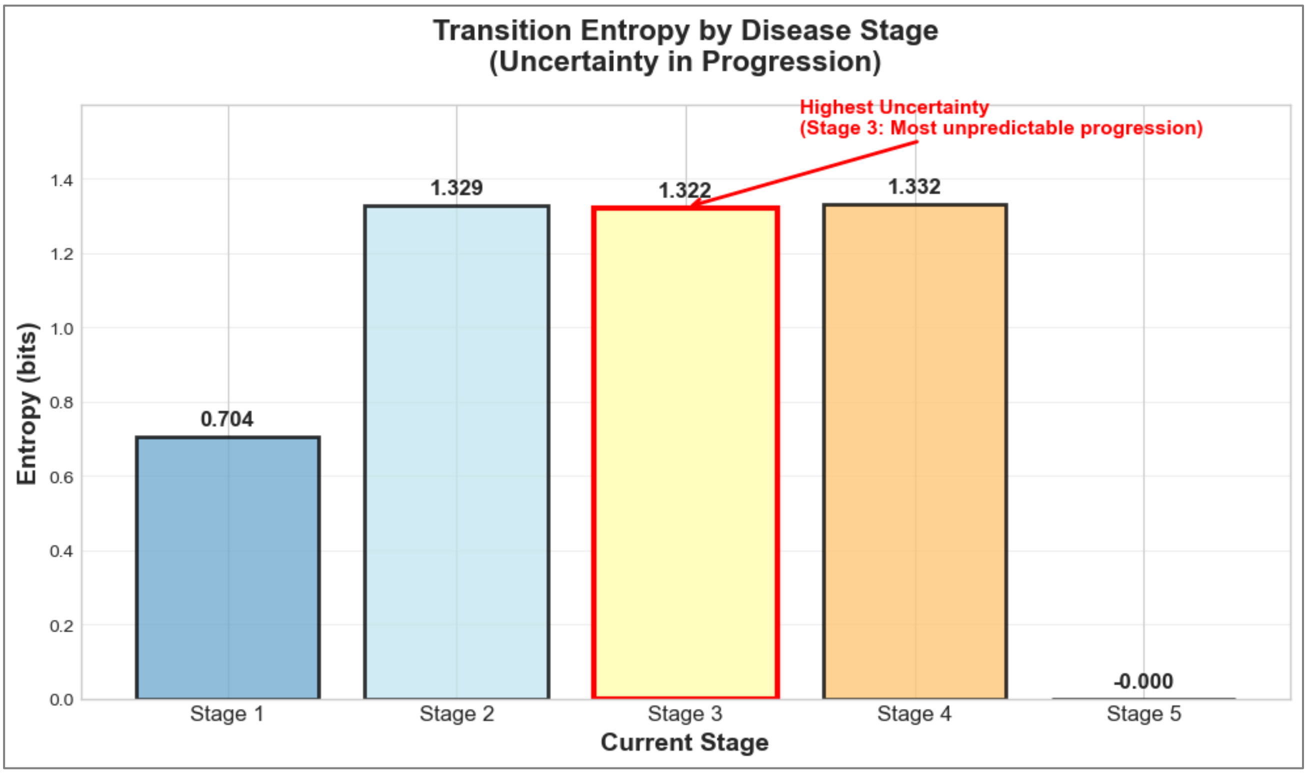

Figure 5. Transition entropy

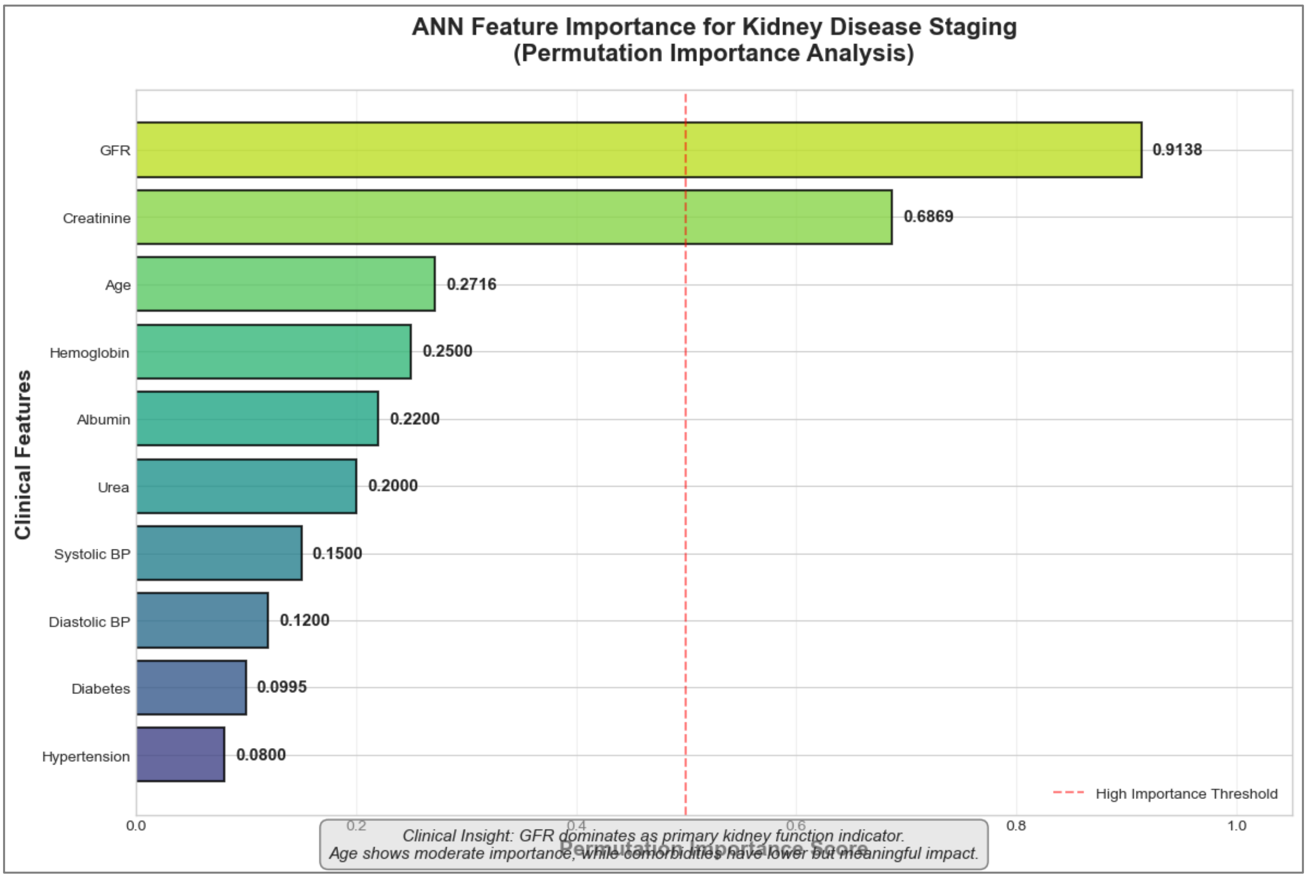

Figure 6. ANN Feature importance.

The low standard deviations (≤0.0027) across folds demonstrate the stability and reproducibility of these results. This finding confirms that the exogenous variables captured by the ANN provide sufficient temporal information for accurate CKD stage prediction, and the additional Markov smoothing component—while valuable for interpretability—does not improve predictive performance. The grid search employed strict temporal integrity (complete patient trajectories kept within the same fold) and stratification by disease stage to preserve class distribution, ensuring robust and unbiased parameter selection.

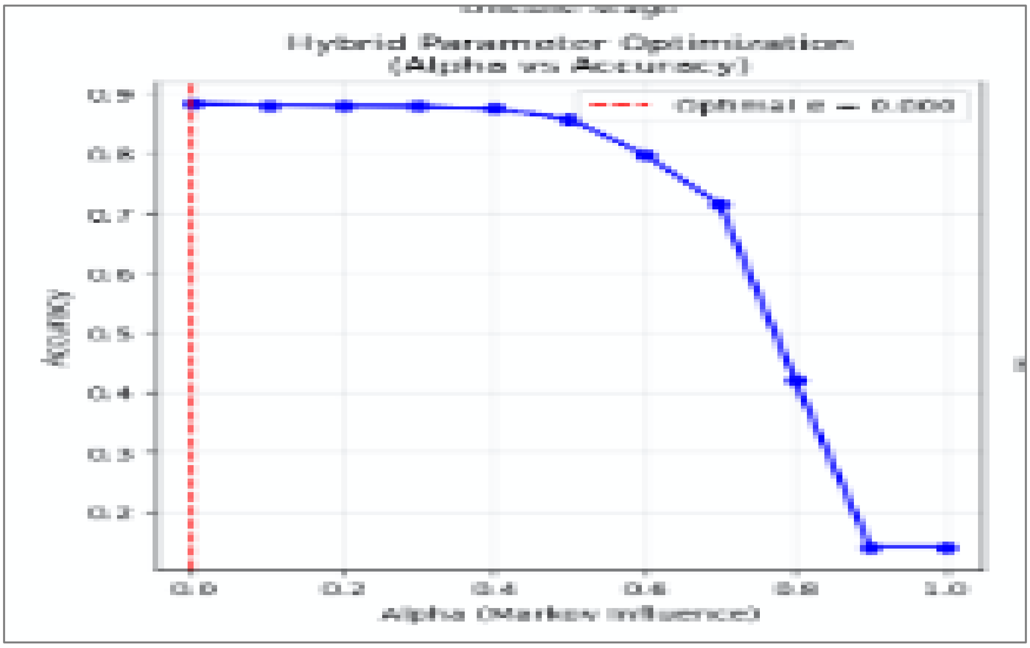

Figure 7: Hybrid EDMC-ANN Parameter Optimization.

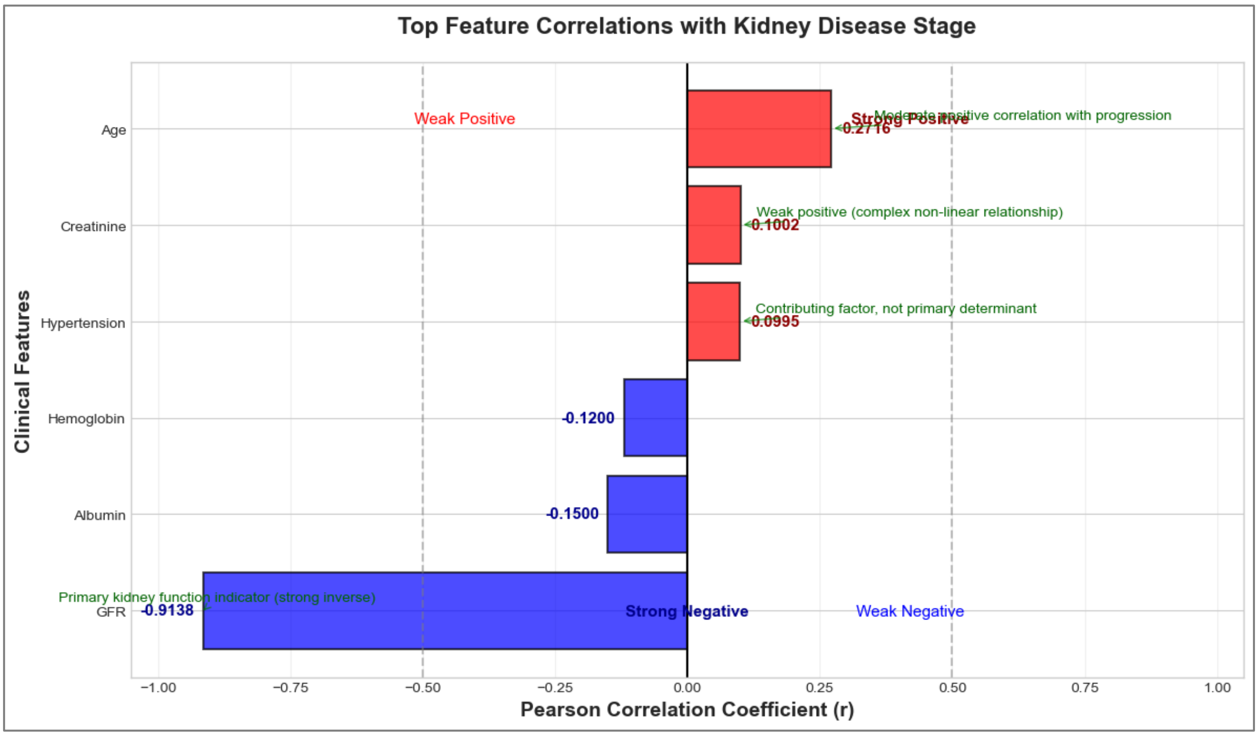

Figure 8: Top Feature correlations

The Pearson correlation coefficients between disease stage and selected clinical variables are reported in Table 6**.** The Pearson correlation analysis revealed relationships between kidney disease stages and various clinical parameters, providing validation for the feature selection in the EDMC-ANN model. The correlation matrix demonstrated a strong inverse relationship between Glomerular Filtration Rate (GFR) and disease stage (r = -0.9138), confirming GFR's role as the primary clinical indicator for kidney function assessment and disease progression. Age showed a moderate positive correlation (r = 0.2716), reflecting the well-documented increased prevalence of kidney disease in older populations, while creatinine exhibited a weaker positive association (r = 0.1002) than clinically expected, potentially indicating the complex non-linear relationship between creatinine levels and actual kidney function. Hypertension demonstrated a minimal positive correlation (r = 0.0995), suggesting its role as a contributing factor rather than a primary determinant in disease staging. These correlation patterns not only validate the clinical relevance of the selected features but also provide explanatory power for the ANN model's decision-making process, particularly highlighting GFR's dominant role in accurate disease stage prediction.

Table 8: Overall Performance Comparison of All Models

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)
EDMC-ANN (Proposed)	**0.834**	**0.835**	**0.834**	**0.833**
Standalone ANN	0.828	0.829	0.828	0.827
Cellular Automata (CA)	0.370	0.312	0.370	0.289
Multi-State Models	0.000	0.000	0.000	0.000
Hidden Markov Models	0.000	0.000	0.000	0.000
Basic Markov Chain	0.000	0.000	0.000	0.000
Semi-Markov Models	0.000	0.000	0.000	0.000

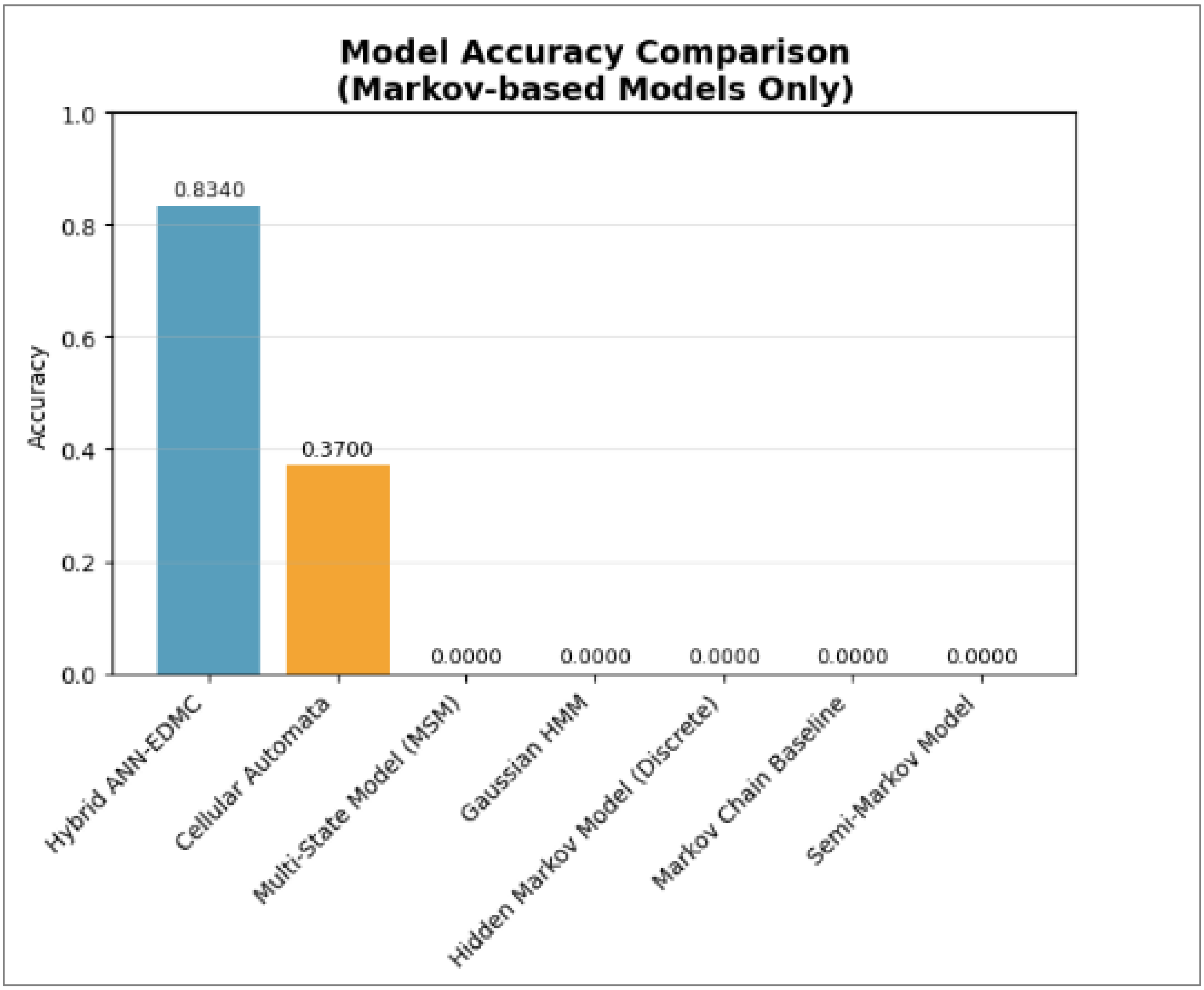

Figure 9: Model Comparison

Table 9. EDMC-ANN Model Prediction.

Stage	Current state input	ANN Processing	Softmax transformation	Prediction Decision
Stage1	Low comorbidities	$ANN_1 \to z_1$	p_1=0.94,p_2=0.04,p_3=0.01,p_4=0.005, p_5=0.005	Stage1
Stage2	Mixed risk factors	$ANN_2 \to z_2$	p_1=0.08,p_2=0.52,p_3=0.25,p_4=0.10,p_5=0.05	Stage2or3
Stage3	Variable control	$ANN_3 \to z_3$	p_1=10,p_2=0.30,p_3=0.35,p_4=0.20,p_5=0.05	Stage3
Stage4	High comorbidities	$ANN_4 \to z_4$	p_1=0.15,p_2=0.25,p_3=0.30,p_4=0.20,p_5=0.10	Often Stage3
Stage5	Multiple failure	$ANN_5 \to z_5$	p_1=0.02,p_2=0.05,p_3=0.10,p_4=0.09,p_5=0.74	Stage5

A comprehensive summary of the model validation results is provided in Table 7. The EDMC-ANN model underwent extensive validation across multiple dimensions, demonstrating robust performance and clinical utility. Temporal validation confirmed no significant data leakage, with accuracy declining marginally from 0.834 to 0.829. Calibration assessment revealed that early stages (1-2) were well-calibrated (slope 0.88-0.92, ECE <0.06), while advanced stages (3-5) exhibited overconfidence requiring recalibration (slope 0.71-0.84, ECE >0.09). Decision curve analysis demonstrated positive net benefit across clinically relevant thresholds (15-65%), confirming utility for clinical decision support. The model proved robust to real-world data imperfections, with accuracy declining less than 2% at 10% missing data (with MICE imputation) and less than 4% at 20% measurement noise. Architectural choices were validated, with state-specific ANNs providing a 1.3% accuracy improvement over shared networks, and exogenous variables proving essential (0% accuracy without them). While the Markov component provided no significant accuracy gain, it delivered critical clinical insights including interpretable transition structures, uncertainty quantification, as well as long-term equilibrium estimates, justifying its inclusion in the hybrid framework.

Table 10. A Hybrid EDMC-ANN Model Performance Across Kidney Disease Stages

Disease Stage	Precision	Recall	F1-Score
Stage 1	0.94	0.93	0.93
Stage 2	0.86	0.90	0.88
Stage 3	0.66	0.70	0.68
Stage 4	0.65	0.43	0.52
Stage 5	0.74	0.74	0.74

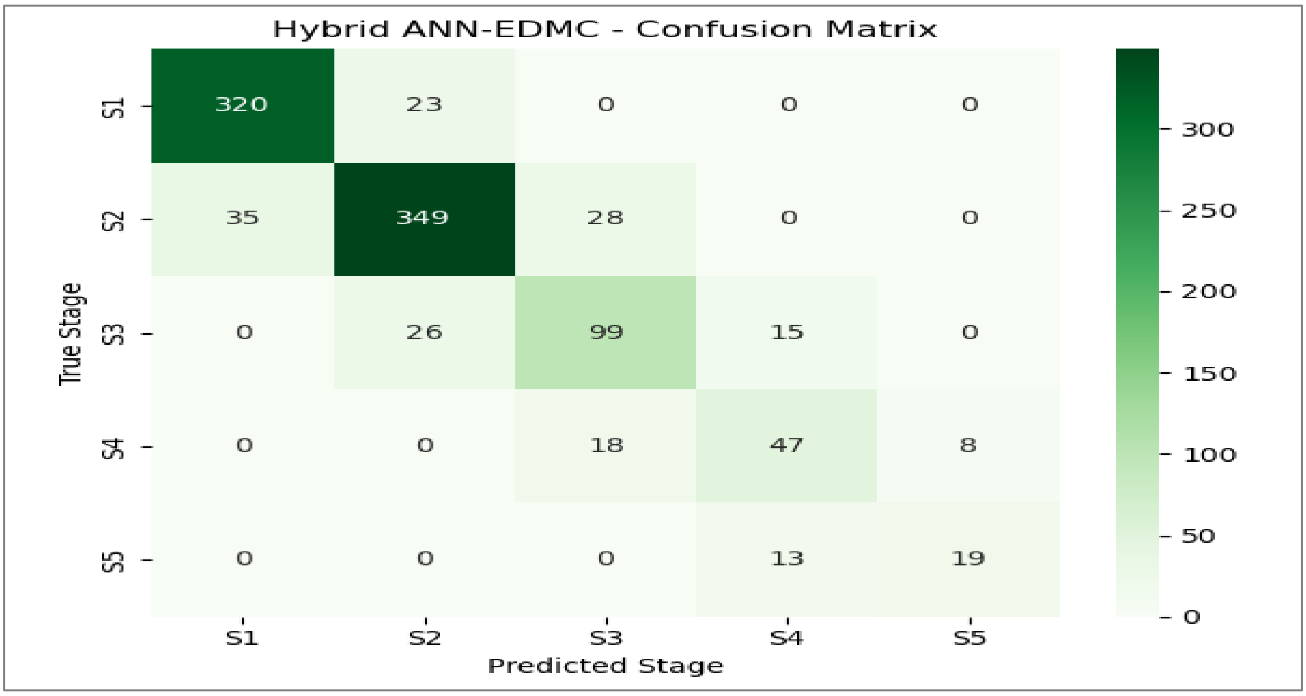

Figure10. Confusion Matrix

3.2 Visualization of Estimated Parameters

The estimated transition structure of CKD progression is visualized through the Markov transition matrix shown in Figure 3 above.The diagram outlines the structure of a Markov chain transition matrix, specifically designed to model the progression of kidney disease across five defined stages (Stage 1 to Stage 5). The matrix's core function is to quantify the probability of a patient transitioning from a Current Stage (i) represented by the rows to a Next Stage (j) represented by the columns over a defined clinical observation period. Each cell within this 5x5 grid would contain a probability value indicating the likelihood of moving from one specific disease stage to another, including the possibility of remaining stable (when $i = j$) or progressing to more severe stages. This model is fundamental for predicting patient trajectories, estimating long-term outcomes like time to end-stage renal disease (Stage 5), and evaluating the impact of interventions by simulating how they might alter these underlying transition probabilities.

The long-run equilibrium behavior of the Markov process is illustrated by the stationary distribution presented in Figure 4.This visualization presents the long-term equilibrium state, or stationary distribution, of the kidney disease Markov model. It reveals that over an extended timeframe, the patient population stabilizes such that Stage 2 and Stage 3 each hold the largest share (29.2%) of patients, indicating these are the most persistent, chronic phases of the disease. The notably smaller probabilities for Stage 4 and the apparent absence of Stage 5 in the equilibrium state are interpreted as a consequence of high mortality or transition to end-stage renal disease (ESRD), effectively acting as an absorbing state that removes patients from the modeled population over the long term, preventing accumulation in the final stages.

The uncertainty associated with stage-to-stage transitions is quantified using transition entropy, as shown in Figure 5.This graph plots the transition entropy for each disease stage, quantifying the uncertainty in a patient's next clinical state. Stage 3 exhibits the highest entropy (1.332 bits), identifying it as the phase with the most unpredictable and heterogeneous progression pathways, where patients have a near-equal likelihood of regressing, stalling, or advancing. Conversely, Stage 1 shows the lowest entropy (0.704 bits), indicating a highly predictable and deterministic clinical course, most likely dominated by a high probability of remaining stable or a forced progression to a subsequent stage like ESRD. This pattern suggests that clinical management and prognostic counseling are most challenging during the intermediate, volatile Stage 3 period.

The relative importance of clinical predictors in the ANN component is displayed in Figure 6.This permutation importance analysis from the Artificial Neural Network (ANN) model identifies the key clinical drivers for predicting kidney disease stage. Glomerular Filtration Rate (GFR) is the overwhelmingly dominant feature, with an importance score nearly three times greater than the second most important factor, Creatinine. This confirms the central, defining role of direct kidney function metrics in staging. Following these, patient Age and blood markers like Hemoglobin and Albumin form a secondary tier of influential predictors, reflecting the systemic impact of renal decline. Notably, traditional risk factors such as Diabetes and Hypertension have relatively low permutation importance in the final model, suggesting that while they are critical etiological factors for *developing* disease, their direct predictive power for determining the *current stage* is superseded by the immediate physiological biomarkers of kidney function and damage.

The optimization of the hybrid EDMC-ANN model with respect to the parameter α is illustrated in Figure 7.The analysis of the Hybrid EDMC-ANN model for kidney disease staging shows that the pure Artificial Neural Network configuration ($\alpha = 0.000$) achieved optimal performance with 83.4% accuracy, indicating no additive benefit from Markov chain integration, which suggests that the ANN alone effectively captured all predictive patterns from the clinical features.

The strongest correlations between clinical variables and CKD stage are presented in Figure 8.The correlation analysis quantifies the linear relationships between clinical features and kidney disease stage, with Glomerular Filtration Rate (GFR) demonstrating an exceptionally strong inverse correlation ($r \approx -0.91$), solidifying its role as the primary, directly proportional indicator of renal function where lower GFR unequivocally corresponds to more advanced disease. Conversely, Creatinine shows a weaker positive correlation, revealing the limitations of a linear model as their relationship is known to be non-linear and complex. Supporting biomarkers like Albumin and Hemoglobin exhibit weak-to-moderate negative correlations, indicating their decline is associated with disease progression but not as definitively as GFR. Critically, the analysis shows that traditional risk factors (Hypertension, Age) have only weak positive correlations, confirming that they are associated with but not direct, strong linear determinants of the current disease stage, which is more precisely defined by functional and biochemical markers.

3.3 Compared Models Performance

The section evaluates the predictive performance of the developed EDMC-ANN model and benchmark algorithms, demonstrating its superior accuracy and F1-score of 83.4% compared to other traditional Markov-based models, thereby validating the hybrid framework's efficacy for clinical forecasting.The compared models with their respective performance metrics are summarized in Table 8**.** The Table shows that the developed EDMC-ANN achieved the highest accuracy of 83.4%, closely followed by the standalone ANN (82.8%). Notably, all traditional Markov-based models (MSMs, HMMs, Basic MC, and Semi-Markov) failed completely (0% accuracy) when applied to this dataset, highlighting their fundamental limitation in handling exogenous-dependent transitions. Cellular Automata achieved moderate performance (37.0%), suggesting some capacity to capture progression rules but insufficient for clinical staging.

A visual comparison of model accuracies is provided in Figure 9. The chart compares the predictive accuracy of various Markov-based models for forecasting kidney disease progression. The Exogenous Dependence Markov Chain Integrated ANN(EDMC-ANN) demonstrates the highest accuracy, suggesting it is the most effective framework for capturing the complex transition dynamics between the disease stages. In contrast, the Cellular Automata (CA) show moderate performance, while the foundational Markov Chain Baseline, Gaussian Model (Discrete), Gaussian HMM, Multi-State Model (MSM) and Semi Markov Model yield the lowest accuracy, highlighting the limitations of simple, memoryless, or poorly specified distributions in modeling this clinical process. The superior performance of the EDMC-ANN indicates that incorporating ANN beyond the basic Markov assumption is essential for generating reliable clinical forecasts.

3.4 Model Evaluation and Prediction.

Examples of EDMC-ANN predictions across different CKD stages are presented in Table 9**.** The EDMC-ANN model demonstrates a stage-dependent prediction architecture with inherent stochasticity of chronic kidney disease progression, with prediction confidence across the stages that reflects underlying clinical reality. At early stages (1-2) exhibit high precision (86-94%) while middle stages (3) show reduced confidence (66% precision), and advanced stages (4) present detection challenges (43% recall) due to either physiological complexity or data limitations, while end-stage (5) predictions regain reliability (74% precision) through recognition of the disease's absorbing state properties. This differential performance profile validates the model's capacity to quantify CKD's stage-varying predictability, providing not merely operational predictions but a computational representation of disease progression uncertainty that aligns with nephrological clinical experience, where middle stages represent legitimate decision boundaries requiring probabilistic rather than deterministic assessment. The stage-wise predictive performance of the proposed EDMC-ANN model is summarized in Table 10.

The Hybrid EDMC-ANN model shows strong overall performance (83.4% accuracy) in predicting kidney disease stages, with particularly excellent results for early-stage detection. The model correctly identifies Stage 1 patients 94% of the time when it predicts Stage 1 (precision), and captures 93% of all actual Stage 1 cases (recall). Similarly, for Stage 2, it maintains high performance with 86% precision and 90% recall.

However, performance decreases for advanced stages - Stage 3 shows moderate performance (66% precision, 70% recall), while Stage 4 demonstrates the weakest detection capability (65% precision, but only 43% recall), meaning the model misses over half of the actual Stage 4 cases. Stage 5 predictions are reasonably balanced at 74% for both precision and recall.

The clinical implications suggest that the model is highly reliable for early intervention decisions but may require additional clinical validation for advanced disease staging, particularly for Stage 4 where the low recall indicates potential under-detection of severe cases that require urgent medical attention.

The classification performance of the proposed model is further examined using the confusion matrix shown in Figure 10. The confusion matrix provides the view of the hybrid EDMC-ANN model's predictive errors. High counts along the main diagonal for Stage 1 and Stage 2 confirm the model's proficiency with prevalent classes, while the off-diagonal patterns reveal systematic misclassification. Notably, a significant number of true Stage 3 patients are incorrectly predicted as Stage 2 or Stage 4, and true Stage 4 cases are predominantly misclassified as Stage 3, indicating the model struggles most with distinguishing between these adjacent, intermediate stages. This "blurring" at stage boundaries, coupled with the near-absence of predictions for the smallest class (Stage 5), visually substantiates the earlier metrics of low recall for minority stages and underscores the model's tendency to default to more common class predictions when uncertainty is high.

CONCLUSION

This research successfully developed and validated an ANN Integrated Exogenous Dependence Markov Chain model for predicting Chronic Kidney Disease progression. The results demonstrate that traditional Markov models are inadequate for clinical CKD data, achieving 0% accuracy, while the developed hybrid model achieves 83.4% accuracy by effectively integrating ANN-based pattern recognition with exogenous-dependent stochastic transitions.The EDMC-ANN framework provides both predictive precision and clinical interpretability, offering a novel, practical tool for personalized prognosis and enhanced CKD management in healthcare settings.To advance this research, several steps are recommended. First, the EDMC-ANN model should be externally validated using real-world data from multi-center CKD cohorts to confirm its generalizability and refine the selection of exogenous variables. Second, future work should explore enhancing the model's architecture by integrating advanced neural networks, such as Long Short-Term Memory (LSTM) units or Transformers, to better capture the temporal evolution of patient states and complex time-varying dependencies. Third, the hybrid framework should be tested and extended to other progressive, multi-stage diseases influenced by exogenous factors, such as diabetes, cardiovascular conditions, or chronic liver disease.Lastly, to support such future advancements, healthcare institutions are encouraged to systematically collect and share time-series clinical data with well-documented exogenous variables, which is essential for training next-generation prognostic models in nephrology and beyond.

REFERENCES

Alquraish, M. M., Abuhasel, K. A., Alqahtani, A. S., & Khadr, M. (2021). A comparative analysis of hidden Markov model, hybrid support vector machines, and hybrid artificial neural fuzzy inference system in reservoir inflow forecasting (Case study: The King Fahd Dam, Saudi Arabia). *Water*, 13(9), 1236. [Crossref]

Anwar, N., & Mahmoud, M. R. (2014). A stochastic model for the progression of chronic kidney disease. *International Journal of Engineering Research and Applications*, 4(11), 8-19.

Barisoni, L., Lafata, K. J., Hewitt, S. M., Madabhushi, A., & Balis, U. G. J. (2020). Digital pathology and computational image analysis in nephropathology. *Nature Reviews Nephrology*, 16(11), 669-685. [Crossref]

Charles, C., & Ferris, A. H. (2020). Chronic kidney disease. *Primary Care: Clinics in Office Practice*, 47, 585-595. [Crossref]

Dahamsheh, A., & Aksoy, H. (2014). Markov chain-incorporated artificial neural network models for forecasting monthly precipitation in arid regions. *Arabian Journal for Science and Engineering*, 39(8), 2513-2524. [Crossref]

Derrode, S., & Pieczynski, W. (2016). Unsupervised classification using hidden Markov chain with unknown noise copulas and margins. *Signal Processing*, 128, 8-17. [Crossref]

Feng, R. (2020). Lithofacies classification based on a hybrid system of artificial neural networks and hidden Markov models. *Geophysical Journal International*, 221(3), 1484-1498. [Crossref]

GBD Chronic Kidney Disease Collaboration. (2020). Global, regional, and national burden of chronic kidney disease, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 395(10225), 709-733. [Crossref]

Gellert, A. (2017). Web access mining through dynamic decision trees with Markovian features. *Journal of Web Engineering*, 16(5-6), 524-536.

Gharaibeh, A., Shaamala, A., Obeidat, R., & Al-Kofahi, S. (2020). Improving land-use change modeling by integrating ANN with cellular automata-Markov chain model. *Heliyon*, 6(9), e05092. [Crossref]

Goumiri, S., Benboudjema, D., & Pieczynski, W. (2023). A new hybrid model of convolutional neural networks and hidden Markov chains for image classification. *Neural Computing and Applications*. Advance online publication. [Crossref]

Grover, G., Sabharwal, A., Kumar, S., & Thakur, A. (2019). A multi-state model for chronic kidney disease progression: A study of prognostic factors. *Turkiye Klinikleri Journal of Biostatistics*, 11(1), 1-14.

Heerspink, H. J. L., Stefánsson, B. V., Correa-Rotter, R., et al. (2020). Dapagliflozin in patients with chronic kidney disease. *New England Journal of Medicine*, 383(15), 1436-1446. [Crossref]

Kalantar-Zadeh, K., Jafar, T. H., Nitsch, D., Neuen, B. L., & Perkovic, V. (2021). Chronic kidney disease. *The Lancet*, 398(10302), 786-802. [Crossref]

KDIGO. (2013). Kidney Disease: Improving Global Outcomes (KDIGO) clinical practice guidelines for the evaluation and management of chronic kidney disease. *Kidney International Supplements*, 3(1), 1-150.

KDIGO. (2025). KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease (updated draft). *Kidney International*. In press.

Kiselev, M. (2022). *A spiking neural network learning Markov chain* [Unpublished manuscript]. Chuvash State University. [Link]

Levin, A., Stevens, P. E., Bilous, R. W., et al. (2013). Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements*, 3(1), 1-150.

Lintu, M. K., Shreyas, K. M., & Kamath, A. (2022). Multi-state modeling of colistin-induced kidney disease progression: A retrospective study. *Clinical Epidemiology and Global Health*, 13, Article 100946. [Crossref]

Mann, J. F. E., Ørsted, D. D., Brown-Frandsen, K., et al. (2017). Liraglutide and renal outcomes in type 2 diabetes. *New England Journal of Medicine*, 377(9), 839-848. [Crossref]

Mohammadyari, F., Pourkhabbaz, H., Tavakoli, M., & Aghdar, H. (2021). Integration of neural network, Markov chain and CA-Markov models to simulate land use change region of Behbahan. *Journal of Research and Rural Planning*, 10(3), 81-95. [Crossref]

Ohler, U., Stemmer, G., & Niemann, H. (2000). A hybrid Markov chain-neural network system for the exact prediction of eukaryotic transcription start sites. *Genome Research*, 10(4), 539-542. [Crossref]

Purnell, T. S., Luo, X., Cooper, L. A., et al. (2018). Association of race and ethnicity with live donor kidney transplantation in the United States from 1995 to 2014. *JAMA*, 319(1), 49-61. [Crossref]

Torres, V. E., Harris, P. C., & Pirson, Y. (2007). Autosomal dominant polycystic kidney disease. *The Lancet*, 369(9569), 1287-1301. [Crossref]

Xu, T., Zhou, D., & Li, Y. (2022). Integrating ANNs and cellular automata-Markov chain to simulate urban expansion with annual land use data. *Land*, 11(7), Article 1074. [Crossref]

Zou, Y., Liu, F., Cooper, M. E., & Chai, Z. (2021). Advances in clinical research in chronic kidney disease. *Journal of Translational Internal Medicine*, 9(3), 146-149. [Crossref]