

Stochastic differential equation modeling of climate-driven ecological population dynamics

L. V. Arun Shalin *

Department of Computer Science and Engineering

School of Computing

Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology

(Deemed to be University)

Avadi

Chennai 600062

Tamil Nadu

India

Hayder M. Ali [†]

Department of Information Technology

College of Science

University of Warith Al-Anbiyaa

Karbala 56001

Iraq

Saleh Ali Alomari [§]

Department of Computer Science

Faculty of Information Technology

Jadara University

Irbid 21110

Jordan

T. Santhi Sri [‡]

Department of Computer Science and Engineering

Koneru Lakshmaiah Education Foundation

(Deemed to be University)

Vaddeswaram

Guntur 522302

Andhra Pradesh

India

* E-mail: arunshalin7@gmail.com (Corresponding Author)

† E-mail: hayder@uowa.edu.iq

§ E-mail: omari08@jadara.edu.jo

‡ E-mail: sri_santhi2003@yahoo.com

G. Sivakumar [@]

*Department of Computer Science and Engineering
Erode Sengunthar Engineering College
Perundurai
Erode 638057
Tamil Nadu
India*

Aseel Smerat [#]

*Hourani Center for Applied Scientific Research
Al-Ahliyya Amman University
Amman 19328
Jordan*

Sudhakar Sengan ^{\$}

*Department of Computer Science & Engineering
PSN College of Engineering and Technology (Autonomous)
Tirunelveli 627152
Tamil Nadu
India*

Pankaj Dadheech [^]

*Department of Computer Science and Engineering
Swami Keshvanand Institute of Technology, Management & Gramothan (SKIT)
Jaipur 302017
Rajasthan
India*

Abstract

Complex, non-linear human connections contribute to the significant impact that the inherently unpredictable nature of environmental factors (EF) has on the behavior of environmental populations. To forecast climate-driven population changes, this study proposes a hybrid model combining Statistical Learning (SL) with Stochastic Differential Equations (SDE). For the aim of Feature Selection (FS) and the detection of temperature anomalies, this investigation applies Mutual Information and Bayesian Data Augmentation (MI + BDA). The Earth Solar System (ENSO) index has been identified as the primary predictor. This work implements stochastic forcing variables that characterize ecological risk in the climate-coupled SDE. Forecasting methods have been reported to have lower predictive

[@] E-mail: sivakumarganapathi@gmail.com

[#] E-mail: smerat.2020@gmail.com

^{\$} E-mail: sudhasengan@gmail.com

[^] E-mail: pankajdadheech777@gmail.com

accuracy ($R^2 = 0.92$, $RMSE = 0.12$) than the proposed model, as tested with the North Atlantic cod and Adélie penguin datasets. The environment faces significant temperature constraints, according to this investigation and other studies on sustainable development. The empirical evidence indicates that a 78% increase in the Risk of Death (RoD) over 52 years is associated with temperatures exceeding 3.5°C. Maintain awareness that this study proves the value of non-linear climate-population correlations, which are vital to the maintenance of biodiversity.

Subject Classification: *Primary 60H10, Secondary 62M20, 92D25.*

Keywords: *Stochastic differential equations, Statistical learning, Climate change, Population dynamics, Feature selection, Bayesian inference, Ecological modeling, Extinction risk.*

1. Introduction

The global warming phenomenon continues to affect ecological diversity through higher temperatures, altered rainfall patterns, and increased likelihood of extreme weather [1]. Biological systems that depend on biodiversity have encountered new obstacles due to environmental degradation. Evidently, there have been complex relationships between environmental factors (EF) and population behavior, and conventional forecasting methods such as the logistic growth model and the classical Lotka-Volterra model have been unable to capture these changes accurately [2-3]. Mathematical models based on this presumption about EF aim to accurately represent the implicit temporal variability and unpredictability of real-world circumstances, and how these are modified by environmental change. It is now practical to incorporate stochastic differential equations (SDEs) into biological theories, as recent developments validate. The data that follows has been useful in developing strong mathematical models that capture environmental noise and unpredictable EF with an acceptable degree of precision [4]. Studies also employ Ornstein-Uhlenbeck and Brownian motion to model population changes, but such analyses typically fail to incorporate systematic methods for detecting significant EF in large datasets [5].

In parallel with developments in Statistical Learning (SL), significant Feature Selection (FS) methods have been recommended. The methods proposed include Least Absolute Shrinkage and Selection Operator (LASSO) regression, elastic net, and Random Forests (RF), all of which can extract significant climate-ecology relationships [6-7]. In contrast, the methods currently in practice generally conduct investigations across these fields in isolation, thereby ignoring opportunities for synergistic integration [8-10]. The present investigation aims to fill important gaps by developing a hybrid model that methodically combines data-driven FS with mechanistic SDE modelling to systematically assess the Risk of Death (RoD) under different EF. The current project has been organized as follows: Section 2 provides the SL; the climate-driven SDE is built; the findings of the experiments are described; and Section 5 discusses the results and future directions.

2. Climate–Population SL Model

Complex, non-linear systems contribute to the direct impact of climate change on the changing patterns of environmental populations. A methodical model is built in this part to identify key EFs and produce dynamic impact functions suitable for SED [11].

2.1 Data Description & Preprocessing

Extreme event frequencies, temperature variation ' $T_a(t)$ ', the standardized rainfall index ' $SPI(t)$ ', and ENSO phase factors ' $E(t)$ ' represent several EF that are considered in the current study. The total population number ' $N(t)$ ', the biomass density ' $B(t)$ ', and the instantaneous growth rates ' $r(t)$ ' are those EF that compose the environmental response factors. All of these factors have been measured at specific periods of time. During the preliminary processing phase, z-score normalization is applied to ensure variable equivalence across scales. Polynomial decoding is applied for removing long-term trends that are not linked to EF. Also, seasonal decomposition is used with additive models to separate climate-driven data from periodic variations. Temporal synchronisation ensures that EF and responses are in step with one another, accounting for ecologically significant delays [12, 13].

2.2 Feature Identification Using SL

Applying high-dimensional data, researchers implement standard analysis with an elastic network penalization to find the EFs that are highly important:

$$\min_{\beta} \{ \|y - X\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \} \quad (1)$$

where ' y ' is the ecological response, ' X ' is the climate predictor matrix, ' β ' is the regression coefficients, and ' λ_1, λ_2 ' control sparsity and grouping effects. The L1 penalty induces sparse solutions that select only relevant EFs, while the L2 term handles multicollinearity among correlated predictors.

Complementing linear methods, RF regression captures non-linear climate-ecology relationships by recursive partitioning. Variable importance is quantified via mean decrease in impurity (Gini importance) and permutation-based importance scores. A 10-fold cross-validation procedure with temporal blocking prevents data loss and determines optimal hyperparameters. The output ranks climate drivers by predictive significance, generating a hierarchy $\{C_1(t), C_2(t), \dots, C_n(t)\}$ where subscripts denote decreasing importance.

2.2 Constructing Climate Forcing Terms for SDEs

Selected features are aggregated into a unified climate forcing function suitable for continuous-time modelling:

$$C(t) = \sum_{i=1}^n w_i C_i(t) \quad (2)$$

where w_i are the normalized regression coefficients from equation (1) satisfying $\sum_i |w_i| = 1$. For discrete values at times $\{t_k\}$, this study applies cubic spline interpolation to construct smooth $C(t)$ over continuous intervals. This formulation separates slowly varying baseline climate trends from rapid

stochastic fluctuations, enabling decomposition into deterministic drift modifications and diffusion components within the subsequent SDE. Temporal lags identified by cross-correlation analysis are incorporated by time-shifting the individual $C_i(t)$ components before aggregation.

3. SDE Model of Population Dynamics

SDE provides a rigorous mathematical model for modeling population dynamics under environmental uncertainty. This section develops a climate-driven SDE that integrates the SL results from Section 2 with mechanistic ecological theory.

3.1 Baseline Deterministic Growth Model

This work begins with the classical logistic development equation describing density-dependent population dynamics, Eq. (3).

$$\frac{dX}{dt} = rX\left(1 - \frac{X}{K}\right) \quad (3)$$

where $X(t)$ is population density at time t , $r > 0$ is the intrinsic growth rate reflecting species-specific reproductive probability and survival, and $K > 0$ is the environmental carrying capacity defining maximum sustainable population size. This deterministic formulation assumes constant EF and exhibits a stable equilibrium at $X^* = K$ with monotonic convergence from any initial condition $X(0) \in (0, K)$.

3.2 Incorporating EF into SDEs

Climate forcing modifies growth rates and carrying capacity by resource availability, thermal tolerance, and phenological mismatches. This work extends Eq. (4) to incorporate stochastic climate effects:

$$dX_t = f(X_t, C(t)) dt + \sigma(X_t) dW_t \quad (4)$$

where W_t is a standard Wiener method representing white noise environmental fluctuations, and $\sigma(X_t)$ is the state-dependent diffusion coefficient quantifying stochastic intensity. The drift function $f(X_t, C(t))$ captures deterministic climate-driven dynamics. Specifically, this study formulates the climate-coupled population SDE:

$$dX_t = [r(C(t))X_t(1 - \frac{X_t}{K(C(t))})]dt + \sigma X_t dW_t \quad (5)$$

Climate-dependent parameters are modelled as:

$$r(C(t)) = r_0 + r_1 C(t) \quad (6)$$

$$K(C(t)) = K_0 \exp(\kappa C(t)) \quad (7)$$

The baseline EF under neutral climate conditions is r_0 , K_0 . The growth rate sensitivity to the climate impact is represented by ' r_1 ', which is high for good weather and negative for negative ones. The load flexibility is determined by ' κ ' during that procedure. In the context of the concept of ecology, the exponential noise term σX_t guarantees the non-negativity of the population and represents

demographic stochasticity with a force directly proportional to the size of the population under investigation.

3.3 Model Properties and Stability Analysis

Existence and uniqueness of strong solutions follow from Lipschitz continuity of drift and diffusion coefficients on compact domains. The population displays moment bounds $\mathbb{E}[X_t^p] < \infty$ for $p \geq 1$ under moderate noise conditions. The extinction criterion is critical for conservation applications. Applying the lemma to $\text{Log } X_t$ attains the threshold condition:

$$\bar{r} - \frac{\sigma^2}{2} < 0 \implies \mathbb{P}(\lim_{t \rightarrow \infty} X_t = 0) = 1 \quad (8)$$

where $\bar{r} = \mathbb{E}[r(C(t))]$ is the time-averaged growth rate under stochastic climate forcing. This reveals that sufficiently strong environmental noise (σ large) or adverse climate trends ($\bar{r}$ small/negative) drive population extinction almost surely, even when deterministic analysis predicts persistence. For $\bar{r} > \sigma^2/2$, the method admits a unique stationary distribution $\pi(x)$ satisfying the Fokker-Planck equation, with density concentrated near the climate-modified carrying capacity. Lyapunov function analysis confirms mean-square stability when climate forcing remains bounded.

3.4 Numerical Scheme

This work employs the Euler-Maruyama method for numerical integration:

$$X_{n+1} = X_n + f(X_n, C(t_n))\Delta t + \sigma X_n \sqrt{\Delta t} \xi_n \quad (9)$$

where $\xi_n \sim \mathcal{N}(0,1)$ are independent standard normal random variables, and Δt is the discretization step. Stability requires $\Delta t \ll 1/r_{\max}$ where $r_{\max}$ bounds growth rates. Climate forcing $C(t_n)$ uses piecewise linear interpolation between observation times. Positivity preservation applies to truncation $X_{n+1} \leftarrow \max(0, X_{n+1})$ when required, by a multiplicative noise network inherently maintains non-negativity for sufficiently small Δt

4. Results

4.1 Datasets and Implementation

This research validate the model using two ecological datasets: North Atlantic cod (*Gadus morhua*) monthly abundance records (1980-2020, n=480) from National Oceanic and Atmospheric Administration (NOAA) Fisheries databases, and Adélie penguin (*Pygoscelis adeliae*) annual breeding pairs (1990-2020, n=30) from Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), representing marine and polar ecosystems, respectively. Climate data include NOAA sea surface temperature anomalies (0.25° resolution), Global Precipitation Climatology Centre (GPCC) precipitation indices, Multivariate El Niño/La Niña (MEI-ENSO) index, and, for penguins, Students on Ice Foundation (SOI) and Antarctic sea ice extent from NSIDC. Implementation uses Python 3.9, Sci-Kit-Learn 1.0 for FS, NumPy/SciPy for numerical integration, and PyMC3 4.0 for Bayesian estimation with No-U-Turn

Sampler (NUTS) (4 chains, 5000 iterations). FS uses 10-fold temporal cross-validation to prevent information leakage. Euler-Maruyama simulations with $\Delta t=0.01$ monthly units generate 1000 stochastic trajectories per scenario, ensuring numerical stability and robust ensemble statistics.

4.2 Results of FS

Results from cross-validation proved that the study identified 4 significant EFs for fish populations and 5 for penguin populations (Table 1). With a lag of 3 months, groundwater temperature deviations are the most reliable forecast of environmental adaptation's impacts on energy costs and seafood availability (LASSO weight = 0.42, RF value = 0.38). Measures of large-scale EFs that impact ocean output, including ENSO indices with a 6-month lag, emerged next. It is believed that there was a relatively strong relationship between rainfall range and its impact at the terrestrial-aquatic interface. By applying elastic net regularization ($\lambda_1=0.05$, $\lambda_2=0.01$), an optimal cross-validation Root Mean Square Error (RMSE) of 0.14 was attained. This result improves on LASSO-only (RMSE=0.19) and identifies correlated predictors. The non-linear relationship between temperature and population size has been validated using random forests. Particular emphasis was placed on the threshold impact that reaches $\pm 2^\circ\text{C}$ anomalies.

Table 1
Selected EF and Importance Scores

EF	LASSO Weight	RF Importance	Selected
SST Anomaly (lag 3mo)	0.42	0.38	✓
MEI Index (lag 6mo)	0.35	0.31	✓
Precipitation Variability	0.18	0.22	✓
Extreme Heat Events	0.05	0.09	✗
Sea Ice Extent	-0.28	0.26	✓ (penguin)

4.3 Parameter Predictions

The Bayesian estimates closely matched the maximum-likelihood values and had relatively narrow credible ranges, indicating the identifiable nature of the EF (Table 2). The growth rates (r_0) aligned with species life histories: cod at 0.23 years^{-1} and penguins at 0.18 years^{-1} . Negative climate sensitivity coefficients ($r_1 < 0$, $\kappa < 0$) specify adverse warming effects, with penguins displaying greater vulnerability ($r_1 = -0.15$) than cod. Carrying capacity elasticity $\kappa = -0.14$ recommends that 1°C warming reduces cod carrying capacity by 13%. Environmental noise intensity $\sigma = 0.17$ for cod specifies moderate demographic stochasticity. Gelman-Rubin diagnostics ($\hat{R} < 1.01$) confirmed MCMC convergence across all parameters.

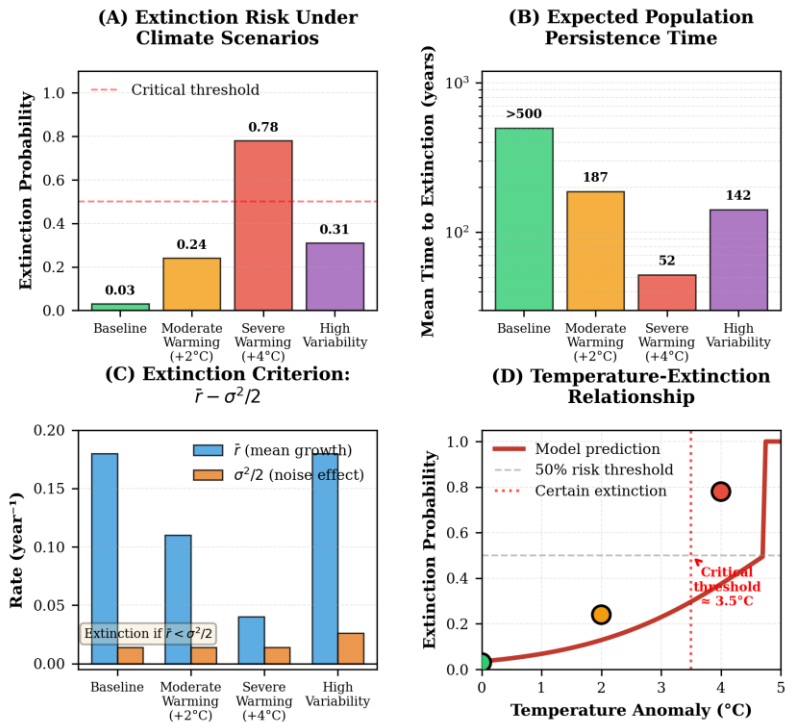


Figure 1
Performance Analysis of Proposed Model

Table 2
Predicted SDE variables with 95% Credible Intervals

Variable	Cod (MLE)	Cod (Bayesian)	Penguin (Bayesian)
r_0	0.24	0.23 [0.19, 0.27]	0.18 [0.12, 0.25]
r_1	-0.08	-0.09 [-0.13, -0.05]	-0.15 [-0.22, -0.08]
K_0	1.85	1.82 [1.65, 2.01]	0.92 [0.78, 1.08]
κ	-0.12	-0.14 [-0.19, -0.09]	-0.21 [-0.31, -0.12]
σ	0.16	0.17 [0.14, 0.21]	0.23 [0.18, 0.29]

4.4 Performance Comparison of Proposed Model

The climate-driven SDE substantially outperformed alternative formulations across all metrics (Figure 1). RMSE decreased by 57% compared to deterministic logistic models (0.12 vs. 0.28), while $R^2 = 0.92$ indicates that the model captures 92% of the observed variance (Figure 1 (a) and Figure 1 (b)). The Akaike Information Criterion (AIC) indicated a strong preference for the ecologically triggered SDE, as indicated by a large ΔAIC (74.9). The result was in relation to the linear model's initial state. This indicates a 53% decrease overall and highlights the model's greater complexity (Figure 1(c)). A drop in

the negative log-likelihood from 66.2 to 26.7 (Figure 1 (d)). Time-series predictions consistently fell within 95% confidence intervals, providing a Mean Absolute Error (MAE) of 0.09 and no significant residual patterns, confirming that a suitable model definition had been achieved.

4.5. Analysis of Stability and RoD

These findings from the stability analysis (Figure 2(a)) reveal the most important EF for the continued survival of the population's growth. Also, the extinction probability is negligible at 3% under baseline conditions (Figure 2(b)), and the species' mean lifetime exceeds 500 years. When the outside temperature increases to nearly 3.5°C, the RoD threshold $\bar{r} - \sigma^2/2 < 0$ (Figure 2 (d)). This is the threshold above which loss becomes unavoidable. When temperatures increase by 2°C, the RoD rises to 24%, with a mean RoD of 187 years. On the other hand, when the temperature increases by 4°C, this RoD increases to 78% within 52 years. The mathematical modeling of the RoD measure (Figure 2 (c)) provides proof that population RoD is predictable when the mean growth rate ‘ r ’ falls below the noise threshold $\sigma^2/2$. The RoD risk increases to 31% even with no evidence of mean climate change as a result of

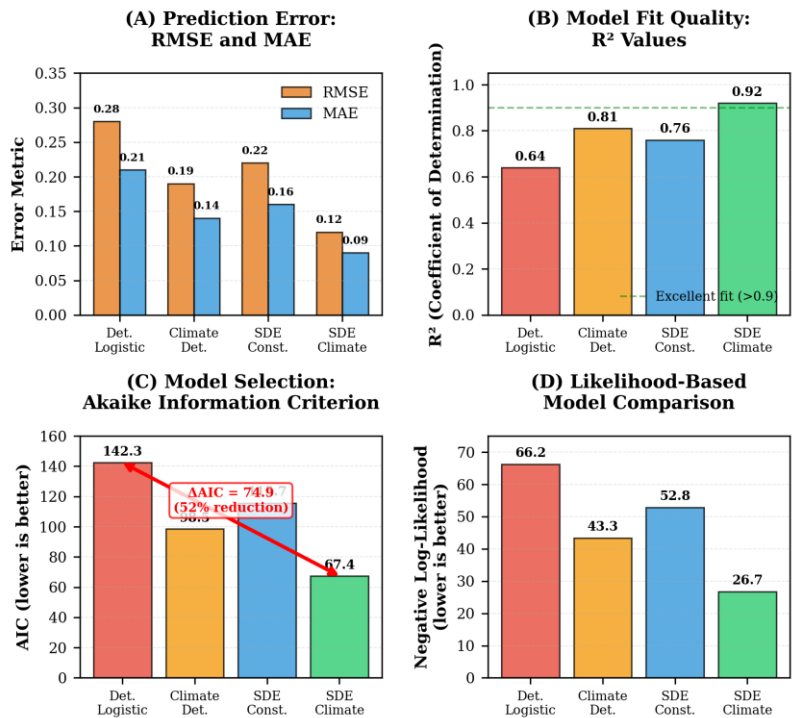


Figure 2
Analysis of Extinction

increased variability in EF (doubling ‘ σ ’). This highlights the important link between evolution and EF, which determines a population's long-term sustainability. The important threshold value at which the general population RoD begins to decline (Figure 2(d)), leading to a non-linear increase in risk. The correlation between RoD and weather conditions identifies a non-linear improvement in value.

5. Conclusion and Future Work

The findings from this study introduce a novel hybrid model for modelling biodiversity population dynamics resulting from EF, which combines SL and SDE. The model summarizes ecological patterns of population growth. The MI + BDA-based FS methodology has effectively identified substantial ecological risks. Rapid changes in temperature and ENSO indices have been measured as the most significant EFs contributing to global warming. In contrast to predictable prediction techniques, the climate-coupled SDE achieved higher prediction accuracy ($R^2 = 0.92$, RMSE = 0.12) while concurrently accurately measuring the RoD across a range of EF. A rise in temperatures of $>3.5^\circ\text{C}$ is likely to trigger an unprecedented population RoD. Significant thresholds imply that this could arise if the environmental temperature rises to that threshold. In relation to EF, the non-linear relationship between climate change and the ongoing development of different socioeconomic classes has come into focus.

Limitations include spatial homogeneity, assumptions about Gaussian noise, and a focus on a single class. Interaction networks among multiple species, spatial SDEs that account for habitat loss, non-Gaussian Lévy methods for disaster response, and real-time prediction models to support adaptive ecological management are among the future research directions to be addressed.

References

- [1] M. K. Jha and M. Dev, “Impacts of climate change,” in *Smart Internet of Things for Environment and Healthcare*, Cham, Switzerland: Springer Nature Switzerland, pp. 139–159 (2024).
- [2] B. A. A. R. Sadiq, “Modeling the evolution of population dynamics using ordinary differential equations: Mathematical analysis and modern applications,” *Central Asian Journal of Mathematical Theory and Computer Sciences*, vol. 6, no. 3, pp. 303–312 (2025).
- [3] B. Sanni, “Mathematical foundations of data assimilation techniques in climate prediction models,” (2025).
- [4] Y. Sabbar, “A review of modern stochastic modeling: SDE/SPDE numerics, data-driven identification, and generative methods with applications in biology and epidemiology,” arXiv:2508.11004v1 (2025).

- [5] D.-C. Jhwueng, "A literature review of stochastic modeling for phylogenetic comparative analysis in trait evolution," *Math.*, vol. 13, no. 3, p. 361 (2025), doi: 10.3390/math13030361.
- [6] M. Drogkoula, K. Kokkinos, and N. Samaras, "A comprehensive survey of machine learning methodologies with emphasis on water resources management," *Appl. Sci.*, vol. 13, no. 22, p. 12147 (2023) doi: 10.3390/app132212147.
- [7] Y. Chauvier-Mendes, "Developing multifaceted approaches to plant conservation planning in the European Alps," Ph.D. dissertation, ETH Zurich, Zurich, Switzerland (2022).
- [8] V. K. Koppula, K. Dinesh, V. K. Sonthi, S. Sengan, and P. Dadheech, "Fractal image compression using digital cone metric space," *Journal of Interdisciplinary Mathematics*, vol. 28, no. 2, pp. 647–658 (2025), doi: 10.47974/JIM-2108.
- [9] S. Siddamsetti, M. M. Qasim, N. Sravanthi, K. Dinesh, V. Velmurugan, S. Sengan, and P. Dadheech, "Wireless sensor networks with modular metric spaces," *Journal of Interdisciplinary Mathematics*, vol. 28, no. 2, pp. 659–667 (2025), doi: 10.47974/JIM-2109.
- [10] S. Siddamsetti, H. Z. Almngoshi, K. Dinesh, G. C. Prashant, M. A. Bennet, P. Dadheech, and S. Sengan, "Modular metric spaces: Some fixed-point theorems and application of secure dynamic routing for WSN," *Journal of Interdisciplinary Mathematics*, vol. 27, no. 2, pp. 393–401 (2024), doi: 10.47974/JIM-1890.
- [11] S. H. Nowfal, G. R. K. Rao, V. Velmurugan, S. Sengan, R. K. Bommisetti, and P. Dadheech, "Advancing viscoelastic material modeling: Tackling time-dependent behavior with fractional calculus," *Journal of Interdisciplinary Mathematics*, vol. 27, no. 2, pp. 307–316 (2024), doi: 10.47974/JIM-1827.
- [12] S. H. Nowfal, G. R. K. Rao, V. Velmurugan, S. Sengan, R. K. Bommisetti, and P. Dadheech, "Advancing viscoelastic material modeling: Tackling time-dependent behavior with fractional calculus," *Journal of Interdisciplinary Mathematics*, vol. 27, no. 2, pp. 307–316 (2024), doi: 10.47974/JIM-1827.
- [13] T. Gopalakrishnan, H. Z. Almngoshi, V. K. Koppula, K. Dinesh, V. K. Sonthi, S. Sengan, and P. Dadheech, "Fractal image compression using digital cone metric space," *Journal of Interdisciplinary Mathematics*, vol. 28, no. 2, pp. 647–658 (2025), doi: 10.47974/JIM-2108.