ISSN: 0019-5693

INDIAN JOURNAL
OF
THEORETICAL PHYSICS

VOLUME 71 **NOS. 3 & 4**

JULY, 2023 – DECEMBER, 2023

K. C. KAR Memorial Special Issue

Published by the
CALCUTTA INSTITUTE OF THEORETICAL PHYSICS
(Formerly, INSTITUTE OF THEORETICAL PHYSICS)
"BIGNAN KUTIR"
4/1, MOHAN BAGAN LANE, KOLKATA-700004

(Peer-reviewed Journal)

ISSN: 0019-5693

INDIAN JOURNAL
OF
THEORETICAL PHYSICS

[Founder President: Late Prof. K. C. Kar, D. Sc.]

VOLUME 71 **NOS. 3 & 4**

JULY, 2023 – DECEMBER, 2023

INDIAN JOURNAL
OF
THEORETICAL PHYSICS

"BIGNAN KUTIR"
4/1, MOHAN BAGAN LANE, KOLKATA-70004, INDIA

SUBSCRIPTION RATE

INDIA : For Library (Institute)

1500.00 for each volume

FOREIGN : $ 350 for each volume

Drafts, Orders, Enquiries & Claim for Non-Receipt of Journal should be sent to:

CALCUTTA INSTITUTE OF THEORETICAL PHYSICS
(Formerly, INSTITUTE OF THEORETICAL PHYSICS)
"BIGNAN KUTIR"
4/1, MOHAN BAGAN LANE, KOLKATA-700004, India

C O N T E N T S

PREFACE

Professor K C Kar was an outstanding theoretical physicist and an eminent and dedicated educationist, who had spent most of his formal working period as a teacher in Presidency College(now known as Presidency University), Kolkata, with some short stints at Scottish Church College, Kolkata and Serampore College and Rajshahi College. This part of his life spanned three decades of twentieth century extending from the early twenties to the mid-fifties. His active life, however extended well beyond that.He converted his residence at 4/1 Mohan Bagan Lane, Kolkata 700 004 to a small working institution which is named "Institute of Theoretical Physics" and initiated the publication of a journal titled "Indian Journal of Theoretical Physics in 1953.The institute that he founded is still functional and is known as "Calcutta Institute of Theoretical Physics". Prof. Kar has been an inspiring figure for generations of aspiring physicists and certainly one of that elite group of Bengali scientists who made science in Bengal competitive with the rest of the world at times which can be described as "difficult".

Kulesh Chandra Kar was very much interested in research and came in contact with Professor C V Raman from whom he received encouragement to carry out research in Indian Musical instruments in 1922.Next he turned his attention towards statistical mechanics and soon mastered over the subject. His rsearches in theoretical physics covered a wide spectrum of subjects like Acoustic Physics, Wave statistics, Nuclear and Particle Physics ,and Theory of Relativity. During his career he published 123 scientific papers in different journals of National and International repute. He also published five books on some special topics of Physics for advanced learners.

The year 2024 marks the 125th Birth Anniversary of Prof.K C Kar. To celebrate this special occasion we have planned to publish this special issue of IJTP as a tribute to Prof. Kar.

In the advent of 21st century Kofi Annan, The Secretary-General of the United Nations, has addressed the usefulness of science in humankind. "Science has contributed immensely to human progress and to the development of modern society. The application of scientific knowledge continues to furnish powerful means for solving many of the challenges facing humanity, from food security to diseases such as AIDS, from pollution to the proliferation of weapons. Recent advances in information technology, genetics, and biotechnology hold extraordinary prospects for individual wellbeing and that of humankind as a whole."

Considering the scientific community's basic concern for human welfare to solve the complex problems related to society & environment the interdisciplinary approach has been the ONLY way out. As such, Science always has been interdisciplinary in nature. Developing novel mathematical theory inspired scientist for its application in every field of science. Application of physical principle in chemistry, biology, and computer science is very well known. While last decades saw the surge of application of principles in physics and mathematics to the highly diverse fields like economics, climate & environmental science, ecology, social science etc. and becoming increasingly attractive for new generation of scientists. To inspire them for exploration in the unmet research fields like Nano systems in Chemistry, Systems Biology approach to medical science, large scale simulation to understand the climate, Ecological balance to recover earth system etc. require large data to generate

and analyse using mathematical & physics rules, so that predictive systems can be robust and our knowledge in such systems become useful and readily applicable.

Importance of Interdisciplinary research has been cited in many recent articles, one of the important facts is how the different subject has influenced each other, as depicted by the figure 1, here (reference: 1). This elaborates that the research impacts reflected by highly cited paper clusters (used in this reference to quantify) statistically significantly and positively associated with interdisciplinarity influence in the subject.

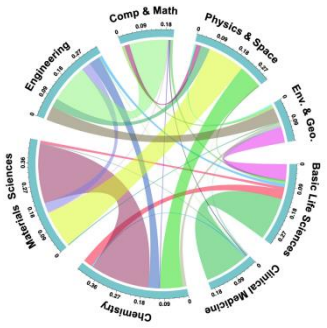


Fig.1. A chord diagram representation of the affinities between research areas. The affinity indices were defined as the time-averaged Jaccard similarity indices and were evaluated between each pair of research areas. They were assigned to each connection between the research areas, represented proportionally by the size of each arc, from which it is evident that the degree of affinity varied considerably for different pairs of the disciplines (see ref1 for details).

As shown in the figure1, Medicine & life sciences are least influenced by physics & mathematics, which provides ample opportunities for the next generations for research contributions, as large data repository and organized data are available freely in these fields (ref2), waiting for mining.

In the present issue, editors have attempted to provide four different articles influencing the field of research in Climate change, Development of new algorithm, Ecological system and in medical science, the application of mathematics & physics as cross disciplinary research articles. Hopefully, this issue will encourage teachers & students to stimulate novel ideas in implementation and development of Interdisciplinary approaches in Science & humanities, which will spread the dream of Prof. Kulesh C Kar widely to scientific communities.

References

1. Keisuke Okamura, Interdisciplinarity revisited: evidence for research impact and dynamism by, https://doi.org/10.1057/s41599-019-0352-4 .
2. Pawan Kumar, Taushif Khan, Indira Ghosh, “Mapping interaction between big spaces; active space from protein structure and available chemical space” as a Book chapter in “Big Data Analytics in Chemoinformatics and Bioinformatics” Edited by Subhash C. Basak & Marjan Vracˇko and published by Elsevier , https://doi.org/10.1016/B978-0-323-85713-0.00029-3 .
3. Susil Kumar Sarkar,“A tribute to Prof. K C Kar”,IJTP Vol.69 Nos 1&2,2021

Untangling Climate's Complexity: Methodological Insights

Alka Yadav[1], Sourish Das[2] and Anirban Chakraborti[1]*

[1] School of Computational and Integrative Sciences,
Jawaharlal Nehru University, New Delhi-110067, India
[2] Chennai Mathematical Institute, Chennai-603103, Tamil Nadu, India

[Abstract: In this article, we review the interdisciplinary techniques (borrowed from physics, mathematics, statistics, machine-learning, etc.) and methodological framework that we have used to understand climate systems, which serve as examples of "complex systems". We believe that this would offer valuable insights to comprehend the complexity of climate variability and pave the way for drafting policies for action against climate change, etc. Our basic aim is to analyse time-series data structures across diverse climate parameters, extract Fourier-transformed features to recognize and model the trends/seasonalities in the climate variables using standard methods like detrended residual series analyses, correlation structures among climate parameters, Granger causal models, and other statistical machine-learning techniques. We cite and briefly explain two case studies: (i) the relationship between the Standardised Precipitation Index (SPI) and specific climate variables including Sea Surface Temperature (SST), El Niño Southern Oscillation (ENSO), and Indian Ocean Dipole (IOD), uncovering temporal shifts in correlations between SPI and these variables, and reveal complex patterns that drive drought and wet climate conditions in South-West Australia; (ii) the complex interactions of North Atlantic Oscillation (NAO) index, with SST and sea ice extent (SIE), potentially arising from positive feedback loops.]

* Corresponding author: anirban@jnu.ac.in

Introduction

"If you think the economy is more important than the environment, try holding your breath while counting your money." – Professor Guy McPherson

The climate system is an excellent example of a "complex system", which is composed of many interconnected and interdependent parts that exhibit "emergent" behaviors and properties not easily predictable from the behavior of individual parts or "sum of its parts"[1–3]. Complex systems are ubiquitous and hence are studied in various domains such as physics, biology, ecology, sociology, economics, environmental science, etc. They exhibit characteristics like: (i) nonlinearity: Small changes in one part of the system can lead to significant and often unpredictable effects throughout the system; (ii) emergence: Novel properties or behaviors emerge at higher levels of organization that are not directly attributable to the individual components of the system; (iii) dynamical behavior and adaptation: systems often exhibit dynamic behaviors such as self-organization, chaos, phase transitions, and they have the ability to adapt and evolve in response to changes in their environment or internal dynamics; (iv) feedback loops: Interactions among system components create feedback loops, where the output of a process feeds back into the system, influencing further interactions, and often leading to catastrophic instabilities[4]. Understanding the dynamics of complex systems, therefore, necessitates a multidisciplinary approach integrating mathematics, physics, statistics, machine learning, and other tools of data science[5]. We must mention here the work of Klaus Hasselmann, a German climate scientist and Nobel Laureate, who made substantial contributions to our understanding of climate dynamics and the development of climate models. In his groundbreaking 1976 work[6], Hasselmann introduced the concept of stochastic climate modeling, incorporating random processes into climate models to elucidate how natural variability and random factors can influence long-term climate trends. He also devised statistical methods to distinguish between natural climate variability and human-induced effects, providing strong evidence of human impact on global warming. Additionally, he played a key role in developing coupled atmosphere-ocean models, crucial for accurately simulating and predicting climate dynamics. In 2021, Hasselmann, alongside Syukuro Manabe and Giorgio Parisi, received the Nobel Prize in Physics[7] for his pioneering contributions to climate modeling and the understanding of complex physical systems. In many social and environmental systems (including climate), we often do not have a clear understanding of the causal relationships between different variables. This makes the understanding of complex climate dynamics significantly more challenging. Therefore, we take the help of Granger causal models[8] and other tools of statistical inference[9,10] to deepen our comprehension of the dynamic interactions among key climate variables, and expand our insights into the intricate mechanisms shaping climate patterns.

Climate change poses a significant challenge globally, affecting ecosystems, economies, well-being of humans as well as livestock[11–14].

> *"Climate change is the greatest threat we face. It's the defining issue of our time, and we have to address it if we want to leave a thriving planet for future generations."* – Katharine Hayhoe, a climate scientist and professor

We hope that understanding of the complex processes behind phenomena of climate change, etc. using insights from multiple disciplines will help us develop effective strategies for mitigation and adaptation. In this respect, we must mention the pioneering works of the economist and Nobel Laureate, William Nordhaus, who developed dynamic and quantitative models (now called integrated assessment models) that described the global interplay between society, the economy, and climate change[15]. The present interdisciplinary research approaches may supplement existing models and their applications in public policy (see e.g., Epilogue in Chakraborti et al.[16]).

Traditionally, people have been building climate models with computer-based simulations of mathematical equations to represent the interactions and processes within the earth's climate system[17–19]. These models have been then used to understand the past climate variations, predict future climate trends, and also assess the potential impacts of climate change. Generally, climate models integrate data on atmospheric dynamics (simulations of the movement of air masses, circulation patterns, and atmospheric processes such as convection, precipitation, and radiation), ocean dynamics (currents, temperature variations, and interactions between the ocean and atmosphere, including El Niño and La Niña), land surface properties (vegetation cover, soil moisture, and land use changes, which influence energy and water exchanges with the atmosphere), and other relevant factors like carbon, nitrogen, and other biogeochemical cycles that play a role in regulating the Earth's climate. Hence, the computer-based climate models provide a physical foundation for the climate change projections and are therefore built to include some of the most comprehensive range of physical, chemical, and biological processes with immense computational complexity; thereby calling for simpler approaches[20]. In the recent past, we too have proposed alternate and simple approaches, based on machine learning and data science, to identify significant statistical relationships among these climate variables and enhance our understanding of climate dynamics[21,22].

In this review article, the basic aim is to introduce the simple methods we have used to analyse time-series data structures across diverse climate parameters, extract Fourier-transformed features to recognize and model the trends/seasonalities in the climate variables. We have used standard methods like detrended residual series analyses, correlation structures among climate parameters, Granger causal models, and other statistical machine-learning techniques. Below, we explain the background and rationale for our two case studies[21,22]:

(1) The relationship between the Standardised Precipitation Index (SPI) and specific climate variables, including Sea Surface Temperature (SST), El Niño Southern Oscillation (ENSO), and Indian Ocean Dipole (IOD)[23–27], uncovering temporal shifts in correlations between SPI and these variables, and reveals complex patterns that drive drought and wet climate

conditions in South-West Australia. Droughts manifest in various forms, with meteorological drought being a critical indicator of extreme climate conditions[28]. Previous studies have explored the dynamics of meteorological droughts and their relationship with climatic factors such as ENSO and IOD cycles[25–27,29–37]. However, there was a necessity for multivariate approaches to understanding drought dynamics, and hence we developed Granger causal models to examine the causal relationships among the variables (SST, NINO 3.4, and IOD) and their collective impact on SPI in South-West Australia, leveraging machine-learning techniques[21].

(2) The complex interactions of the North Atlantic Oscillation (NAO) index, with SST and sea ice extent (SIE), potentially arising from positive feedback loops. We delved into another study[22] focusing on the complex dynamics of climate variables such as the North Atlantic Oscillation (NAO), a key atmospheric pressure index affecting weather patterns across North America and Northern Europe. Past research has highlighted the NAO's substantial impact on cold air outbreaks, storm occurrences, and climate variability in these regions. Previous studies had also underscored the positive feedback loop between melting Arctic SIE and increasing SST, driven by atmospheric new particle formation and growth, accelerating Arctic warming[38]. Recent studies had shown: (i) a significant decrease in SIE in the coming years, intensifying global atmospheric circulation and directly impacting SIE melting[39,40]. (ii) The winter NAO plays an important role in weather variability in northwest Europe, with recent studies highlighting the predictive power of autumnal Arctic sea ice for winter NAO forecasting[41]. (iii) NAO variability accounts for a substantial portion of atmospheric pressure variability and correlates with SST anomalies[42–44]. Besides, climate models have illustrated how multidecadal variations in the NAO induce corresponding fluctuations in Atlantic circulation and Arctic sea ice loss, contributing to hemispheric warming[45]. Hence, we developed a hybrid model to analyse the relationships: (a) the SPI to SST, NINO 3.4, and IOD, (b) the interplay among North Atlantic Oscillation (NAO), SST, and Sea Ice Extent (SIE). Utilizing machine learning algorithms like LASSO, we identified significant Fourier harmonics essential for modeling long-term memory. Additionally, to capture short-term memory, we incorporated lagged estimators such as IOD, SST, and NINO 3.4 within the framework of a Granger causal model. Employing data-driven techniques, we revealed intricate interactions among NAO, sea surface temperature (SST), and sea ice extent (SIE), shedding light on critical instabilities and feedback loops crucial for addressing climate change[38–51]. Our approach had two distinct advantages: (i) It offered a broader perspective for addressing climate change compared to traditional climate forecast models[52]. (ii) While previous studies had focused on specific Arctic regions and seasons, our approach employed statistical machine learning models to provide a comprehensive view of the *entire* North Atlantic region.

Methodology

The primary goal is to analyse and model the climate variables and learn about the intricate interdependences. The climate variables are typically observed and analysed as multivariate time series datasets. Suppose,

$$\mathbf{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{k,t})^T \tag{1}$$

is a vector of different climate variables observed at time point t; where $x_{k,t}$ represents the value of the k^{th} climate variable at a time at t. The superscript T denotes the transpose operation, which converts the row vector into a column vector.

Simple measures of dependences

Autocorrelation function, also known as *serial correlation*[53], is a statistical concept that measures the degree of similarity between a given time series of a climate variable and a lagged version of itself over successive time intervals. It quantifies how much the current value of a climate variable is related to its past values. For example, we may be interested in knowing if the past 30 days' values of SST influence the current value of SST. If this is so, that might help us to predict the possible value of SST 30 days ahead from today. The autocorrelation function (ACF) is described as follows:

$$\text{ACF}(k) = \frac{\text{Cov}(x_t, x_{t-k})}{\sqrt{\text{Var}(x_t) \times \text{Var}(x_{t-k})}} \tag{2}$$

where $\text{ACF}(k)$, is the autocorrelation between the values of the climate variables at time t and those at time $t-k$, and k represents the number of time units (lags) by which the series is shifted. The covariance $\text{Cov}(x_t, x_{t-k})$ is between the values of the climate variables at time t and those at time $t-k$; and $\text{Var}(x_t)$ and $\text{Var}(x_{t-k})$ represent the variance of the climate variables at the time t and those at time $t-k$. A positive ACF value at lag k indicates a positive correlation between the values of the same climate variables separated by k time units.

Cross-Correlation Function (CCF) is a statistical measure that indicates the relationship between values of two different climate variables separated by a time lag[54]. The cross-correlation function (CCF) is defined as:

$$\text{CCF}(k) = \frac{\text{Cov}(x_t, y_{t-k})}{\sqrt{\text{Var}(x_t) \times \text{Var}(y_{t-k})}} \tag{3}$$

where $\text{Cov}(x_t, y_{t-k})$ is the covariance between the variable x_t at time t and the lagged value of variable y_t at time $t-k$. The CCF can take on positive, negative, or zero values. A positive CCF indicates a positive correlation between the two variables at the specified lag, while a negative CCF indicates a negative correlation. A CCF close to zero suggests no significant correlation between the variables at that lag.

Calculating Periodicity using ACF. For examining the enduring memory span of each time series, which indicates the persistence of past observations influencing current values, we utilised the following algorithms:

(i) Calculating the ACF (Eqn. 2) on training data.
(ii) Identifying periods $P_1, P_2, \cdots, P_{m_s}$, where the autocorrelation ρ_m exceeds a threshold s, with $s = \frac{1}{M}\sum_{m=1}^{M} |\rho_m - \hat{\rho}|$; here, ρ_m represents the m^{th} lag autocorrelation, $\hat{\rho}$ is the median of all autocorrelations, and M represents the maximum lag considered in our invastigation study.

To illustrate this algorithm, we take the example of the SPI time series of West South Australia (coordinates: longitude 113.7158 and latitude -26.6969) and plot the ACF (shown later). Utilising a dataset spanning 450 months, ranging from June 1973 to November 2010, we detected three noteworthy periods: 216, 151, and 60 months. These periods suggested that the current SPI value exhibited a notable positive correlation with past SPI values, with periodicities of approximately 5.5 years for the specified location.

Modelling Temporal Structure and Seasonalities

Temporal patterns in data can be effectively modeled to uncover underlying trends and periodic behavior. We introduce a comprehensive model (M_s) to analyse the components contributing to the observed variability:

$$M_s: \; y(t) = \beta_0 + \beta_1 t + \eta(t) + z(t), \tag{4}$$

where (M_s) is the model for a location s time series, $y(t)$ represents the observed variable at time t, β_0 denotes the intercept, β_1 is the trend coefficient, $\eta(t)$ accounts for the periodicity of the process, and $z(t)$ captures short-term memory using the Granger Causal Model (explained below).

In time series analysis, capturing seasonality is important for understanding recurring patterns. Here, we present a model ($\eta(t)$) aimed at quantifying seasonal effects:

$$\eta(t) = \sum_{j=1}^{m_s} \left\{ \sum_{i=1}^{K} \beta_{ji} \sin(i * \omega_j * t) + \sum_{i=1}^{K} \gamma_{ji} \cos(i * \omega_j * t) \right\}. \tag{5}$$

This equation breaks down seasonal variation into Fourier series, where $\omega_j = \frac{2\pi}{P_j}$, $j = 1, 2, \cdots, m_s$, P_j is estimated via ACF, and m_s denotes the number of periods for s^{th} location time series. Then, we typically applied the LASSO (least absolute shrinkage and selection operator) technique[55], a machine learning shrinkage method, to identify the most significant harmonics in the Fourier model. LASSO selects harmonics that significantly reduce error, enhancing model accuracy.

Modelling Spatial Correlation

To estimate $y(t)$ at location s while accounting for spatial correlation, we introduce the estimated value $\hat{y}_s(t)$ for the t^{th} month, employing a spatially correlated Gaussian process model,

$$\bar{y}_s(t) = \Sigma(s, s')[\Sigma(s, s') + \tau^2\mathbf{I}]^{-1}\hat{y}(t), \tag{6}$$

where, $y_s(t)$follows a Gaussian process with a mean of zero and a covariance function defined as $\Sigma(s, s') = \tau^2 \exp -\rho|s - s'|^2$. Note that the covariance matrix $\Sigma(s, s')$ models the spatial correlation between the location s and s' for which $\mathbb{V}ar(\epsilon) = \sigma^2$, and $\Sigma(s, s') = exp\{-\rho|s - s|^2\}$.

Granger-causal model

For the short-term memory or autoregressive structure, denoted as $z(t)$ (mentioned above), we employ the Granger causal model. The Granger causality model serves as a powerful tool for assessing the changing causal dynamics between variables over time[8]. Here, we present hypotheses regarding the influence of one variable on another using this framework.

Null Hypothesis (H_0):

The null hypothesis represents that the variable of interest (Z) is solely dependent on its own historical memory, without any influence from other variables. It can be represented as:

$$z(t) = \beta_0 + \beta_1 z(t-1) + \cdots + \beta_k z(t-k) + \epsilon(t). \tag{7}$$

This equation represents a time series model where $z(t)$ is the variable of interest at time t. The term β_0 is the intercept, and $\beta_1, \beta_2, \ldots, \beta_k$ are coefficients that indicate the influence of past values of $z(t)$ on its current value. The $\epsilon(t)$ term represents the error or residual component of the model at time t, assumed to follow a normal distribution with mean 0 and variance σ^2.

Alternate Hypothesis (H_a):

Under the alternate hypothesis, we propose that the variable of interest ($z(t)$) is not only influenced by its own historical memory but also by the past values of another variable ($x(t)$). This can be represented as follows:

$$\begin{aligned} z(t) \;=\; & \beta_0 + \beta_1 z(t-1) + \cdots + \beta_k z(t-k) \\ & + \gamma_1 x(t-1) + \cdots + \gamma_k x(t-k) + \epsilon(t). \end{aligned} \tag{8}$$

In this equation, the coefficients $\gamma_1, \gamma_2, \ldots, \gamma_k$ represent the influence of past values of the variable $x(t)$ on the current value of $z(t)$. Each γ_i indicates the strength and direction of influence from the corresponding lagged value of $x(t)$. This model allows us to investigate whether the variable of interest ($z(t)$) is affected by the past values of another variable ($x(t)$), thereby examining potential causal relationships between them.

To determine whether the null hypothesis should be rejected, we assess whether all coefficients γ_i in the alternate hypothesis are equal to zero. Specifically, the null hypothesis H_0 states that $\gamma_1 = \cdots = \gamma_k = 0$. To reject this null hypothesis in favor of our alternate hypothesis H_a, we need to ascertain if at least one coefficient

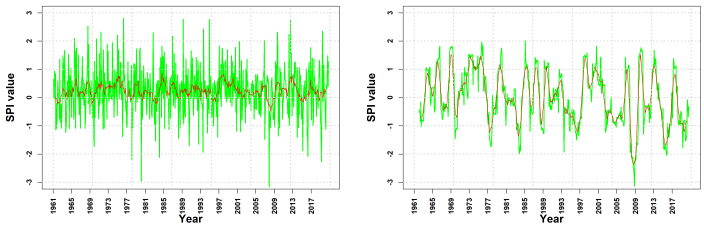


Fig. 1. Time series plots illustrating the Standardised Precipitation Index (SPI) for the south-western region of Australia are depicted for (a) 1-month and (b) 12-month periods, covering a span of 58 years (1961-2018). In both plots, the 12-month moving average is represented by the solid red line. A notable increase in the signal-to-noise ratio is observed in the SPI 12-month series compared to the SPI 1-month series, thereby enhancing the reliability of our analysis[59]. This consistent observation was evident across all 194 monitoring stations, prompting our selection of the SPI 12-month series for further analysis.

γ_i is not equal to zero. This rejection allows us to clearly understand the effect of the variables under consideration.

Two case studies of climate complexity

Complex Dynamics of Drought in South-West Australia

Our first study delves into the complex interactions among climatic variables such as SST, NINO 3.4, and IOD, examining their influence on the SPI in South-West Australia.

Data Description. We utilised four primary climate variables: SPI, SST, NINO 3.4, and IOD indices, with a focus on South-West Australia. Our analysis involved SPI monthly time series data spanning 58 years (1961–2018), obtained from daily precipitation observations across 194 stations in the region, sourced from the Bureau of Meteorology (BOM) website[56]. The datasets for NINO 3.4 and IOD were sourced from the NOAA website[57,58]. Our study area encompasses longitudes 113.72 to 137.12 and latitudes -26.70 to -35.73. Analyses were conducted using monthly averaged SST data from 1982 to 2018[21].

This region holds significant agricultural importance, contributing substantially to Australia's gross agricultural production and grain exports. However, the agriculture sector faces considerable vulnerability due to the unpredictable nature of climate variables, resulting in reduced water supplies and impacting wheat and broadacre livestock production[60]. Climate change exacerbates these challenges, evident from increasing temperature and fluctuating rainfall anomalies[12,14,59,61,62]. Understanding these complexities is crucial, as they severely impact agriculture, water resources, and public health.

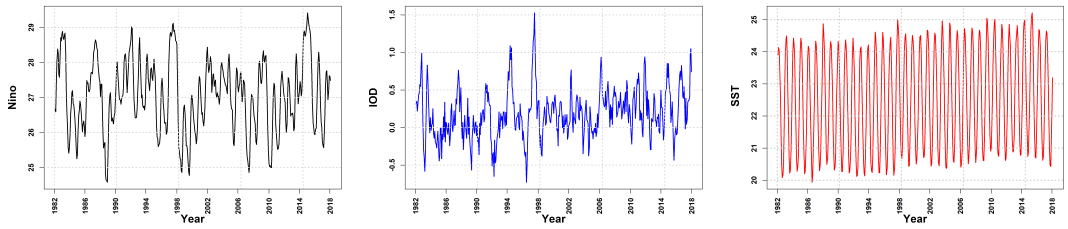


Fig. 2. Time-series plots illustrating (a) El Niño–Southern Oscillation (ENSO) NINO 3.4, (b) Indian Ocean Dipole (IOD), and (c) Sea Surface Temperature (SST) data are presented, covering a period of 37 years (1982-2018) for the south-western region of Australia.

Results and Analyses. Our exploration into the dynamics between the Standardised Precipitation Index (SPI) and climate variables—SST, IOD, and NINO 3.4—across the extensive period from 1982 to 2018 revealed insightful findings. Time series plots in Figure (1) depict the SPI for both one-month and 12-month periods spanning 58 years (1961-2018). The solid red line represents the 12-month moving average, indicating a higher signal-to-noise ratio in the SPI 12-month series compared to the SPI 1-month series. This observation, supported by[59], bolsters the reliability of our analysis. The consistency of this observation across all 194 monitoring stations prompted us to select the SPI 12-month series for our further analysis. Our investigation also unveiled a tendency for SST, Nin 3.4, and IOD to exhibit mean-reverting behavior, as evidenced in Figure (2), where values tended to regress to their means over time. This observation aligns with findings from the Hurst exponent coefficient analysis[63], Table (1) indicates a long memory characteristic in these variables, with values surpassing 0.5.

Figure (3) provides illustrative insight into one of the 194 locations studied. Utilising algorithm , we identified significant periodicities at intervals of 60, 151, and 216 months. For instance, at the specific location (longitude 113.7158 and latitude -26.6969), the SPI displayed a notable positive correlation with past SPI values, recurring at approximately 5.5-year intervals. Such findings contribute valuable insights into the complex interplay between precipitation patterns and key climatic drivers, shedding light on the underlying mechanisms governing long-term climate variability.

Figure (4) is a visual representation of the correlation matrix illustrating SPI among the 194 locations. This matrix underscores the spatial correlation observed across all locations, highlighting the cohesive nature of SPI dynamics throughout the region.

To further explore the dynamics, we analysed cross-correlation functions () among monthly time-series of SPI, SST, IOD, and NINO 3.4 variables, as shown in Figures (5) and (6). With a maximum lag of 120 months, these CCFs reveal mutual influences among climate variables. Figure (5) illustrates the Cross-Correlation Function (CCF) spanning the years 1999 to 2008, revealing the interplay among climate variables (SST, NINO 3.4, and IOD) and their mutual

influence. This interaction results in a dense network, as depicted in Figure (7a). In Figure (6), spanning 2009-2018, we observe that SST weakly couples with SPI (Figure (6a)), while IOD does not directly couple with SPI (Figure (6b)). Conversely, NINO 3.4 couples with SPI (Figure (6c)). Additionally, IOD couples with SST and NINO 3.4, though not reciprocally (Figures (6d) and (6e)). Notably, NINO 3.4 and SST exhibit mutual coupling (Figure (6f)). This comprehensive analysis, spanning 2009–2018, is depicted as a network diagram in Figure (7b).

Our hybrid model (), combining Fourier harmonics for long-term memory and Granger causal modeling () for short-term memory, effectively captured SPI dynamics. Validation through Root Mean Square Error (RMSE) assessments demonstrated (Table (2)) the superiority of models incorporating LASSO selection and Gaussian process (GP) spatial correction (), particularly when considering SST as a covariate[55]. Notably, the incorporation of SST significantly improved SPI estimation accuracy. In Figure (8), we presented the out-of-sample estimates alongside the actual SPI values. Visual examination suggests satisfactory performance of the proposed model (4) for December 2010. However, it's crucial to emphasize that this validation is specific to a single month. To comprehensively evaluate the model's effectiveness, we extended the assessment period from December 2010 to November 2018 (eight years) and calculated the out-of-sample Root Mean Square Error (RMSE) for all 194 locations. The median RMSE results are outlined in Table (2). The first model (Type I) used Fourier series methods on SPI to capture long-term memory, while other models (Type II, III, IV) integrated Nino 3.4, IOD, and SST as covariates, employing various combinations and lags. Our investigation revealed that incorporating the LASSO approach for the Fourier model with spatial correction notably enhanced model performance, with Type III and Type IV models demonstrating superior efficacy. Notably, SPI 12 months exhibited lower RMSE values across all Type II, III, and IV models, indicating enhanced generalizability compared to SPI 1 month. Furthermore, our study identified SST as a significant factor influencing SPI estimation, and its inclusion as a covariate contributed to improved model performance.

Analysis of SPI trends in Figure (9) revealed evolving drought conditions, with inland areas showing mild increasing trends, particularly post-2008. The negative correlation between IOD and SPI until 2008 shifted to a positive correlation thereafter, indicating a reversal in drought conditions[64]. Similarly, NINO 3.4 exhibited a consistent negative relationship with SPI, suggesting wetter conditions in south-west Australia. Rising SST trends further corroborated expectations of increasingly wet conditions in the region[65].

Overall, our findings underscore the complex interplay of climate variables and the evolving nature of drought conditions in south-west Australia, emphasizing the importance of considering long-term trends and spatial correlations in climate modeling and prediction. Specifically, the LASSO selection process within spatial correction models, effectively enhances the accuracy of the proposed models.

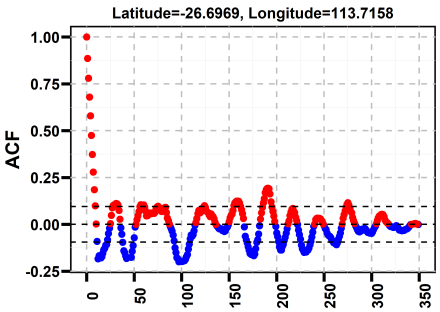


Fig. 3. The autocorrelation plot, with a maximum lag of 400 months, pertains to rainfall data obtained from a specific location at longitude 113.7158 and latitude -26.6969. A dataset spanning 450 months, from June 1973 to November 2010, was utilized to construct the autocorrelation function. Notably, three significant periods of 216, 151, and 60 months were identified. These periods indicate a substantial positive correlation between current and past SPI values, suggesting a periodicity of approximately 5.5 years.

Index	**Hurst Value**
Standard Precipitation Index (SPI)	0.71 (±0.05)
El Niño Southern Oscillation (ENSO) NINO 3.4	0.66 (±0.04)
Indian Ocean Dipole (IOD)	0.69 (±0.06)
Sea Surface Temperature (SST)	0.58 (±0.05)

Table 1. This table displays the Hurst exponent values along with Bootstrap margins of error for all indices, including SPI, IOD, SST, and NINO 3.4. These values collectively suggest the presence of long memory within the system. The table was reproduced from Yadav et al.[21]

North Atlantic Climate Instabilities

Our second study incorporates three distinct datasets: daily mean Arctic Sea Ice Extent (SIE), daily mean Sea Surface Temperature (SST), and daily mean North Atlantic Oscillation (NAO) index.

Data Description. The NAO and SST datasets are sourced from the NOAA website[58,66], while the SIE dataset is obtained from the National Snow and Ice Data Centre's website[67]. These datasets cover various time periods, with NAO data available from 1950, SIE data from 1979, and SST data from 1982. Our analysis spans from January 1982 to September 2019, encompassing 38 years.

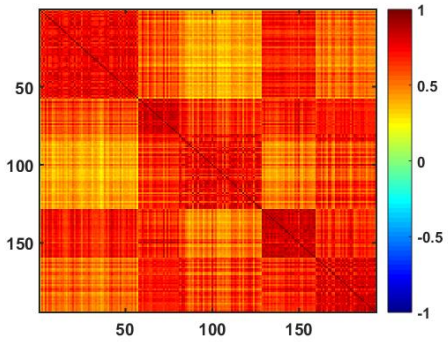


Fig. 4. A visual representation of the correlation matrix depicting SPI-12 months across 194 locations reveals significant spatial correlation across all locations.

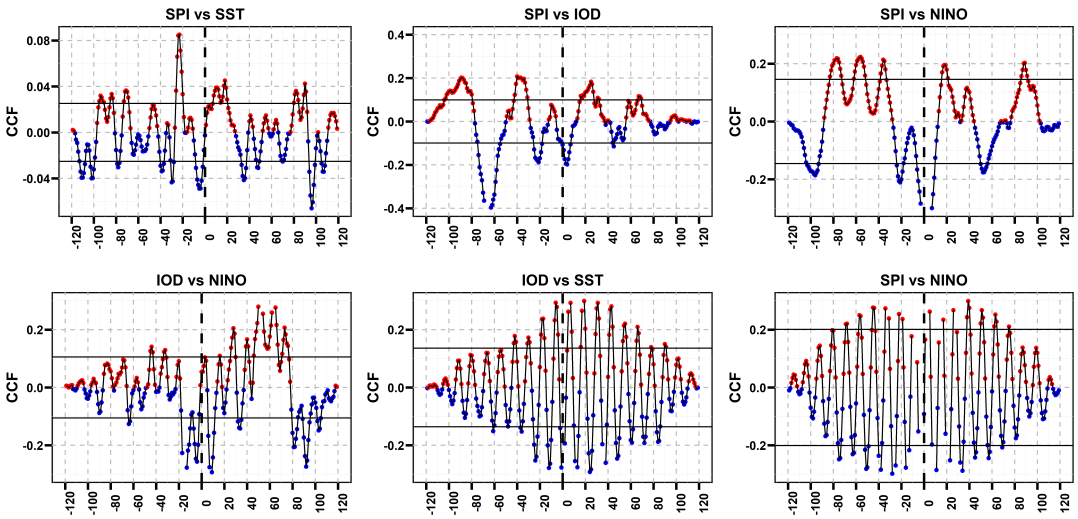


Fig. 5. During the period from 1999 to 2008, cross-correlation function (CCF) plots were generated for the following pairs: (a) Standardized Precipitation Index (SPI) and Sea Surface Temperature (SST), (b) SPI and Indian Ocean Dipole (IOD), (c) SPI and Nino 3.4, (d) IOD and Nino 3.4, (e) IOD and SST, and (f) Nino and SST, with a lag of 120 months. The cross-correlation coefficients indicate significant coupling among these indices (Nino, IOD, and SST). All plots were generated using monthly temporal datasets. The figure was reproduced from Yadav et al.[21]

Results and Analyses. Figure (10a) depicts the NAO plot spanning from 1979 to 2019. NAO, which signifies the difference in sea-level air pressure between the Icelandic Low and the Azores, demonstrates a stationary process with a mean of zero. The time series plot confirms the mean-zero stationary nature of NAO. This observation was further validated through the Augmented Dickey-Fuller test, as evidenced by the small p-value. Additionally, the Auto-correlation function

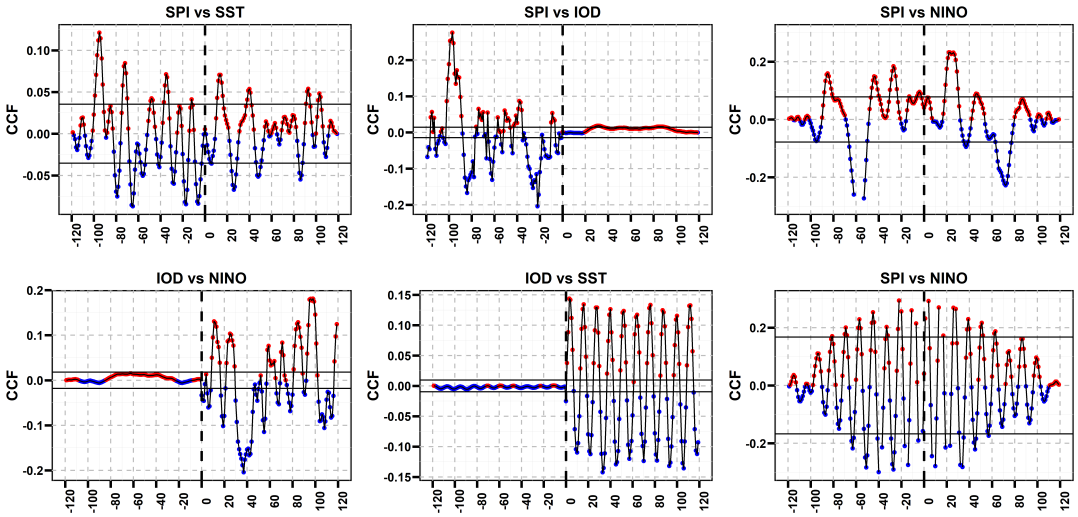


Fig. 6. From 2009 to 2018, cross-correlation function (CCF) plots were generated for the following pairs: (a) SPI and SST, (b) SPI and IOD, (c) SPI and Nino 3.4, (d) IOD and Nino 3.4, (e) IOD and SST, and (f) Nino and SST, with a lag of 120 months. In Figure (a), despite a small CCF value, SPI and SST are observed to exhibit coupling. Figure (b) reveals a one-way cross-correlation between IOD and SPI, and vice versa, albeit statistically insignificant. Figure (c) highlights a significant coupling between Nino 3.4 and SPI. Conversely, Figure (d) indicates that while IOD influences Nino 3.4, the reverse is not observed. Figure (e) suggests a minor effect of IOD on SST, with no reciprocal influence observed. Notably, both SST and Nino 3.4 exhibit significant coupling. The substantial CCF coefficients indicate the significant influence of SST, Nino, and IOD on each other. Plots were generated using monthly temporal datasets. The figure was reproduced from Yadav et al.[21]

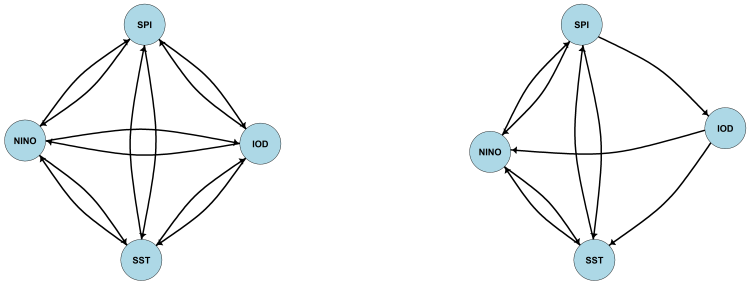
.

(ACF) analysis illustrated in Figure (10b) with a maximum lag of 5000 days (approximately 13 years) reveals the presence of long memory.

Figure(11) represents the residual for SIE and SST obtained from the best fit of Model (4). It is observed that both SIER and SSTR exhibit characteristics of zero-mean stationary processes, similar to the North Atlantic Oscillation (NAO).

The Hurst exponent values detailed in Table (3) collectively underscore the long memory inherent in residuals of SIE (SIER) and residual of SST (SSTR), with values surpassing 0.5. Additionally, the correlation matrix in Table (4) highlights robust correlations, particularly notable between SSTR and NAO, and between SIER and SSTR.

To explore the potential presence of a feedback loop between the SIER, SSTR, and NAO by employed the Granger causality test . Our exploration into Granger causal models uncovers significant causal relationships among NAO, SIER, and



(a) (b)

Fig. 7. The network analysis based on Cross-Correlation plots (5) and (6) for the periods (a) 1999 to 2008 and (b) 2009 to 2018 reveals distinct patterns. In Figure (a), spanning 1999 to 2008, IOD exhibits direct effects on SST, Nino 3.4, and SPI. However, in the subsequent decade (Figure (b), 2009 to 2018), IOD's direct impact on SPI diminishes. Nevertheless, both SST and Nino 3.4 continue to directly influence SPI in both periods, and vice versa. Consequently, IOD indirectly couples with SPI. The significant CCF coefficients indicate substantial mutual influences among these climate variables. The figure was reproduced from Yadav et al.[21]

Model	**Type I**		**Type II**		**Type III**		**Type IV**	
SPI	**1m**	**12m**	**1m**	**12m**	**1m**	**12m**	**1m**	**12m**
Model without Spatial Correction	**0.75**	1.26	0.80	**0.49**	0.80	**0.51**	0.84	**0.38**
Model with Spatial Correction	7.21	**0.96**	6.40	**0.45**	8.66	**0.46**	24.21	**0.37**
Model with LASSO and Spatial Correction	**0.75**	0.77	0.80	**0.35**	0.80	**0.34**	0.84	**0.34**

Table 2. The out-of-sample Root Mean Square Error (RMSE) values are calculated for all four types of models from December 2010 to November 2018, covering a period of 8 years. Here, the target variable is SPI (1 month and 12 months). A brief overview of the model types is as follows: Type I captures long-term memory solely using Fourier Series methods. Type II incorporates NINO 3.4 and IOD as covariates. Type III includes lag values of NINO 3.4 and IOD as covariates. Type IV incorporates NINO 3.4, IOD, and SST as covariates. It is observed that Type III and Type IV, considering LASSO and GP correction, yield the lowest (best) RMSE value of 0.34 for estimating SPI-12 months. The table was reproduced from Yadav et al.[21]

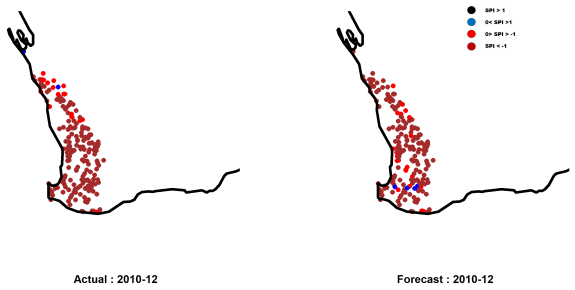


Fig. 8. Plot of the Standardized Precipitation Index (SPI) for December 2010 using training data ranging from June 1973 to November 2010. The SPI values are categorized into four color ranges: Black represents SPI ¿ 1 (Extremely Wet), Blue represents 0 ¡ SPI ¿ 1 (Wet), Red represents 0 ¿ SPI ¿ -1 (Dry), and Brown represents SPI ¡ -1 (Extremely dry). The figure was reproduced from Yadav et al.[21]

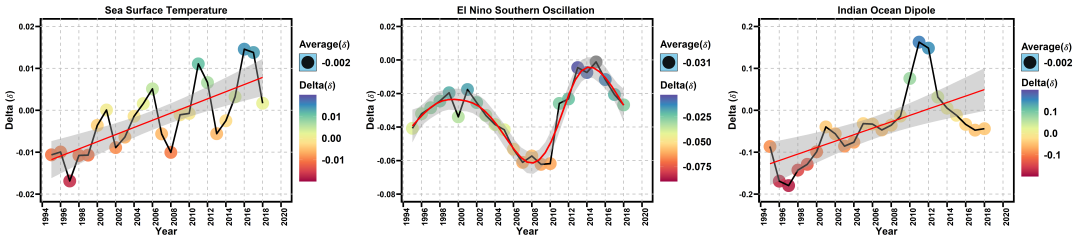


Fig. 9. Plots illustrating the average regression coefficient value (δ) corresponding to NINO 3.4, SST, and IOD are displayed for the period from 1995 to 2018. A red smooth curve represents the three-year moving average of δ. The analysis highlights that NINO 3.4 consistently exerts a significant negative impact on SPI. In contrast, SST exhibits a consistent negative correlation with SPI until 2004. However, between 2005 and 2014, the δ of SST fluctuates between negative and positive correlations. IOD demonstrates a negative correlation with SPI until 2009, followed by a positive range from 2009 to 2013, and then returns to a significantly negative correlation with SPI thereafter. The figure was reproduced from Yadav et al.[21]

SSTR. ANOVA F-tests reject null hypotheses, affirming Granger causality between these variables, as elaborated in Table (5). Furthermore, employing an Akaike information criterion-based model selection process helps identify optimal configurations for Model Equation .

The synthesis of these findings highlights the existence of a reciprocal feedback loop involving NAO, SIER, and SSTR. Moving forward, we aim to demonstrate the affirmative nature of this loop. By emphasizing the skewness of NAO in Table (6), we provide insight into the bootstrap confidence intervals (C.I.) across various time intervals—daily, weekly, and monthly. In a scenario where NAO

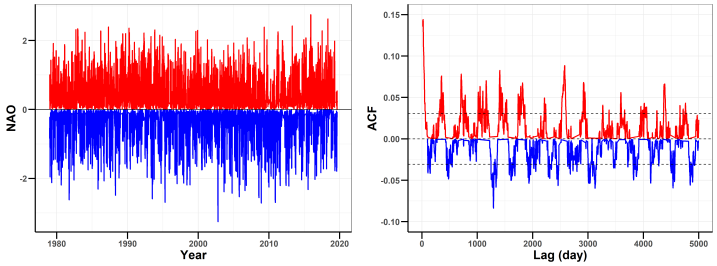


Fig. 10. (a) Plot of the NAO time series over the period of 1979–2019 (b) Autocorrelation plot of NAO with a maximum lag of 5000 days (almost 13 years), indicating the existence of long memory. The figure was reproduced from Yadav et al.[22]

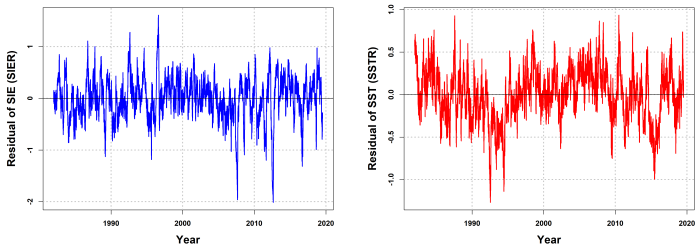


Fig. 11. Time series plots of (a) Residual of SIE (SIER) and (b) Residual of SST (SSTR). The figure was reproduced from Yadav et al.[22]

is stable, a skewness value of zero is anticipated. However, our results reveal a negatively skewed distribution, suggesting a statistically significant departure from stability. Notably, the skewness of NAO in Table (6) indicates a departure from stability, hinting at underlying instability within its dynamics. In Figure (12), we present visualizations of projected and observed trajectories within the test dataset spanning 2010 to 2019. Our modeling approach demonstrates strong performance, as indicated by Root Mean Square Error (RMSE) values of 0.36 for the training set and 0.41 for the test set, underscoring its high generalization capability.

Hurst Exponent	**SIER**	**SSTR**
Simple R/S Hurst estimation	0.77	0.81
Corrected R over S Hurst exponent	0.83	0.89
Empirical Hurst exponent	0.82	0.90
Corrected empirical Hurst exponent	0.81	0.89
Theoretical Hurst exponent	0.53	0.52

Table 3. The Hurst Exponent value of the SIER, and SSTR using different methods. The table was reproduced from Yadav et al.[22]

	NAO	SIER	SSTR
NAO	1.000	0.016 (0.063)	-0.133 ($< 2.2 * 10^{-16}$)
SIER		1.000	-0.173 ($< 2.2 * 10^{-16}$)
SSTR			1.000

Table 4. Over a span of 38 years, from January 1982 to September 2019, the correlation matrix for NAO, SIER, and SSTR is examined. The accompanying P-values enclosed in parentheses provide insights into the significance of the correlations. Notably, the correlation between SSTR and NAO is statistically significant. Similarly, a robust correlation is observed between SIER and SSTR. However, the correlation between NAO and SIER exhibits relatively weak significance. The table was reproduced from Yadav et al.[22]

GC Models	F-value	p-value
SSTR + SIER → NAO	2.31	0.0178
NAO + SIER → SSTR	5.546	2.16×10^{-6}
NAO + SSTR → SIER	7.714	2.27×10^{-10}

Table 5. Table of F-value and p-value of different combinations of Granger Causal Models. Small p-values indicate that there is a feedback loop among NAO, SIER, and SSTR. The table was reproduced from Yadav et al.[22]

Period	Skewness	C.I.
Daily	-.210	[-0.242, -0.169]
Weekly	-.213	[-0.305, -0.107]
Monthly	-.194	[-0.368, -0.005]

Table 6. The presented table highlights the skewness of NAO, along with bootstrap-derived confidence intervals (C.I.), across daily, weekly, and monthly time spans. While an anticipated stable NAO would exhibit a skewness of zero, our observations indicate a negatively skewed distribution. This statistically significant result underscores the presence of instability in NAO. The table was reproduced from Yadav et al.[22]

Our analyses revealed crucial instability driven by positive feedback loops among melting SIE, rising SST, and the North NAO. Key findings included reciprocal Granger causality between SIE and SST, mutual Granger causality between SST and NAO, and an anti-correlation between SST and NAO. This anti-correlation suggested that increasing SST trends may lead to more negative NAO occurrences, a phenomenon supported by the negative skewness of the NAO index across various time scales.

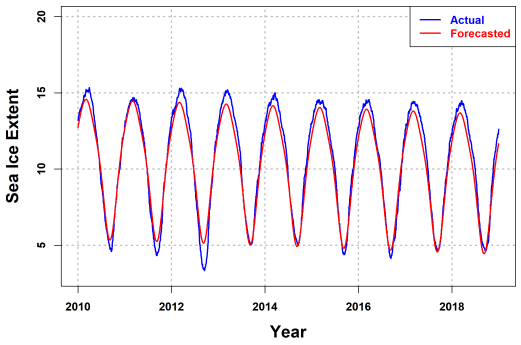


Fig. 12. To evaluate the effectiveness of our proposed model a machine learning assessment was carried out. The training dataset spanned from 1979 to 2009, while the test dataset encompassed the years 2010 to 2019. The figure was reproduced from Yadav et al.[22]

The negative skewness of the NAO index, despite its mean-zero stationary nature, indicated impending critical instability, foreshadowing increased bouts of frigid climates in the North Atlantic region. This study contributed to predicting notable climate transformations by enhancing our understanding of critical instability within complex climate systems. Overall, through the application of statistical machine learning and interdisciplinary methods, as opposed to climate modelling, we tried to enrich the understanding of the dynamic interplay among crucial climate variables and its implications for the entire North Atlantic region.

Outlook

This review study contributes to untangling the complexity (transformations and critical instabilities) of climate systems. Through the application of statistical machine learning and interdisciplinary methods, borrowed from physics, mathematics, statistics, etc. we tried to give a flavour of the alternate approaches to climate modeling.

Climate change presents numerous challenges across various domains, including environmental, social, economic, and geopolitical aspects. Addressing these challenges requires coordinated efforts at local, national, regional, and global levels, including mitigation measures to reduce greenhouse gas emissions, adaptation strategies to build resilience, investments in clean energy and sustainable practices, policy interventions, technological innovations, education, and public awareness campaigns. We hope that these insights would pave the way for drafting policies for action against climate change and addressing its multifaceted challenges. Finally, we accept our limitations as scientists, encapsulated by the famous quote:

> *"I used to think the top environmental problems were biodiversity loss, ecosystems collapse and climate change. I thought that with 30 years of good science we could address those problems. But I was*

wrong. The top environmental problems are selfishness, greed and apathy... and to deal with those we need a spiritual and cultural transformation and we, (Lawyers) and scientists, don't know how to do that." – James Gustave Speth, US Advocate

Acknowledgements

The authors would like to thank K.S. Bakar for some of the collaborative work that has been included in this review article.

References

1. Parisi, G. Complex systems: a physicist's viewpoint. *Physica A: Statistical Mechanics and its Applications* **263**, 557–564 (1999).
2. Vemuri, V. *Modeling of complex systems: an introduction* (Academic Press, 2014).
3. Gell-Mann, M. What is complexity? In *Complexity and industrial clusters: Dynamics and models in theory and practice*, 13–24 (Springer, 2002).
4. Chakraborti, A. *et al.* Emerging spectra characterization of catastrophic instabilities in complex systems. *New Journal of Physics* **22**, 063043 (2020).
5. Chakrabarti, A. S., Bakar, K. S. & Chakraborti, A. *Data Science for Complex Systems* (Cambridge University Press, 2023).
6. Hasselmann, K. Stochastic climate models part i. theory. *Tellus* **28**, 473–485 (1976).
7. The nobel prize in physics 2021. Link to the website: https://www.nobelprize.org/prizes/physics/2021/summary/.
8. Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438 (1969).
9. Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267–288 (1996).
10. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*, vol. XIV, 426 (Springer New York, NY, 2013).
11. What is climate change? Link to the website: https://www.un.org/en/climatechange/what-is-climate-change.
12. Climate change – trends and extremes. Australian Government Bureau of Meteorology, Link to the website: http://www.bom.gov.au/climate/change/.
13. Climate trends in western australia. Department of Primary Industries and Regional Development, Link to the website: https://www.agric.wa.gov.au/climate-change/climate-trends-western-australia.
14. Climate change and broadacre livestock production in western australia. Department of Primary Industries and Regional Development, Link to the website: https://www.agric.wa.gov.au/climate-change/climate-change-and-broadacre-livestock-production-western-australia.
15. Nordhaus, W. D. Reflections on the economics of climate change. *Journal of Economic Perspectives* **7**, 11–25 (1993).
16. Chakraborti, A., Haven, E., Patra, S. & Singh, N. *Quantum decision theory and complexity modelling in economics and public policy* (Springer Nature, 2023).
17. Schneider, S. H. & Dickinson, R. E. Climate modeling. *Reviews of Geophysics* **12**, 447–493 (1974).
18. Hoskins, B. J. Dynamical processes in the atmosphere and the use of models. *Quarterly Journal of the Royal Meteorological Society* **109**, 1–21 (1983).
19. Held, I. M. The gap between simulation and understanding in climate modeling. *Bulletin of the American Meteorological Society* **86**, 1609 – 1614 (2005).
20. When less is more: Opening the door to simpler climate models. Link to the website: https://eos.org/opinions/when-less-is-more-opening-the-door-to-simpler-climate-models.

21. Yadav, A., Das, S., Bakar, K. S. & Chakraborti, A. Understanding the complex dynamics of climate change in south-west australia using machine learning. *Physica A: Statistical Mechanics and its Applications* **627**, 129139 (2023).
22. Yadav, A., Das, S. & Chakraborti, A. Understanding north atlantic climate instabilities and complex interactions using data science (2023). arXiv:2001.10171.
23. Li, W., Fu, R., Negrón-Juárez, R. & Fernandes, K. Observed change of the standardized precipitation index, its potential cause and implications to future climate change in the amazon region. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **363**, 1767–72 (2008).
24. Wu, Renguang, K. & L., J. Analysis of the relationship of us droughts with sst and soil moisture distinguishing the time scale of droughts. *American Meteorological Society* **22**, 4520–4538 (2008).
25. Delage, F. & Power, S. The impact of global warming and the el niño-southern oscillation on seasonal precipitation extremes in australia. *Climate Dynamics* (2020).
26. Wen, N., Liu, Z. & Liu, Y. Direct impact of el niño on east asian summer precipitation in the observation. *Climate Dynamics* **44** (2015).
27. Loughran, T., Pitman, A. & Perkins-Kirkpatrick, S. The el niño–southern oscillation's effect on summer heatwave development mechanisms in australia. *Climate Dynamics* **52** (2018).
28. Heim, R. A review of twentieth–century drought indices used in the united states. *Bulletin of the American Meteorological Society* **83** (2002).
29. McKee, Thomas B., D., Kleist, N. J. & John. The relationship of drought frequency and duration to time scales. *Eighth Conference on Applied Climatology* 17–22 (1993).
30. Karavitis, A., Christos, A., Stavros, T., Demetrios E., A. & George. Application of the standardized precipitation index (spi) in greece. *Water* **3**, 787–805 (2011).
31. Sonmez, F. Kemal, K., Ali Umran Komuscu, E., Ayhan, T. & Ertan. An analysis of spatial and temporal dimension of drought vulnerability in turkey using the standardized precipitation index. *Natural Hazards* **35**, 243–264 (2005).
32. Lee, C., Shen, S., Bailey, B. & North, G. Factor analysis for el nino signals in sea surface temperature and precipitation. *Theoretical and Applied Climatology* **97**, 195–203 (2009).
33. Taschetto, Andréa S., E. & H., M. El niño modoki impacts on australian rainfall. *American Meteorological Society* **22**, 3167–3174 (2009).
34. Koll, R., Gualdi, S., Drbohlav, H.-K. & Navarra, A. Seasonality in the relationship between el nino and indian ocean dipole. *Climate Dynamics* **37**, 221–236 (2011).
35. Ndehedehe, C. E. *et al.* What if the rains do not come? *Journal of Hydrology* **595**, 126040 (2021).
36. Izumo, T. *et al.* Influence of Indian Ocean Dipole and Pacific recharge on following year's El Niño: interdecadal robustness. *Climate Dynamics* **42**, 291–310 (2014).
37. Nicholls & Neville. Sea surface temperatures and australian winter rainfall. *Journal of Climate* **2**, 965–973 (1989).
38. DallÓsto, M. *et al.* Arctic sea ice melt leads to atmospheric new particle formation. *Scientific Reports* **7**, 3318 (2017).
39. Das, P., Lahiri, A. & Das, S. Understanding sea ice melting via functional data analysis. *Current Science* **115**, 920–929 (2018).
40. Cressey, D. Arctic sea ice at record low. *Nature* (2007).
41. Warner, J. L. Arctic sea ice – a driver of the winter nao? *Weather* **73**, 307–310 (2018).
42. Kwok, R. Recent changes in arctic ocean sea ice motion associated with the north atlantic oscillation. *Geophysical Research Letters* **27**, 775–778 (2000). https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1999GL002382.
43. Miettinen, A., Koç, N., Hall, I. R., Godtliebsen, F. & Divine, D. North atlantic sea surface temperatures and their relation to the north atlantic oscillation during the last 230 years. *Climate Dynamics* **36**, 533–543 (2011).
44. Pan, L.-L. Observed positive feedback between the nao and the north atlantic ssta tripole. *Geophysical Research Letters* **32** (2005).

45. Delworth, T. L. *et al.* The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere. *Nature Geoscience* **9**, 509–512 (2016).
46. Jaiser, R., Dethloff, K., Handorf, D. & Cohen, J. Impact of sea ice cover changes on the northern hemisphere atmospheric winter circulation. *Tellus A: Dynamic Meteorology and Oceanography* **64**, 11595 (2012).
47. Becker, G. A. & Pauly, M. Sea surface temperature changes in the north sea and their causes. *ICES Journal of Marine Science* **53**, 887–898 (1996).
48. Slonosky, V. C. & Yiou, P. The north atlantic oscillation and its relationship with near surface temperature. *Geophysical Research Letters* **28**, 807–810 (2001).
49. Dai, G., Mu, M. & Wang, L. The influence of sudden arctic sea-ice thinning on north atlantic oscillation events. *Atmosphere-Ocean* **59**, 39–52 (2021).
50. Horvath, S., Stroeve, J., Rajagopalan, B. & Jahn, A. Arctic sea ice melt onset favored by an atmospheric pressure pattern reminiscent of the north american-eurasian arctic pattern. *Climate Dynamics* **57** (2021).
51. Parkinson, L. C. Recent trend reversals in arctic sea ice extents: Possible connections to the north atlantic oscillation. *Polar Geography* **24**, 1–12 (2000).
52. Kolstad, E. W. & Screen, J. A. Nonstationary relationship between autumn arctic sea ice and the winter north atlantic oscillation. *Geophysical Research Letters* **46**, 7583–7591 (2019).
53. Autocorrelation 2024, april 28. In Wikipedia. https://en.wikipedia.org/wiki/Autocorrelation.
54. Cross-correlation 2024, april 28. In Wikipedia. https://en.wikipedia.org/wiki/Cross-correlation.
55. Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267–288 (1996).
56. Indian ocean dipole. Australian Government Bureau of Meteorology, Link to the website: http://www.bom.gov.au/climate/data/.
57. Global climate observing syatems. Working Group on Surface Pressure, Download Climate Timeseries, Link to the website: https://psl.noaa.gov/gcos_wgsp/Timeseries/.
58. NOAA. Noaa optimum interpolation (oi) sea surface temperature (sst) v2 (2020). Data can be downloaded from: https://psl.noaa.gov/repository/entry/show?entryid=f8d470f4-a072-4c1e-809e-d6116a393818.
59. Lehner, B., Doell, P., Alcamo, J., Henrichs, T. & Kaspar, F. Estimating the impact of global change on flood and drought risks in europe: A continental, integrated analysis. *Climatic Change* **75**, 273–299 (2006).
60. National recovery and resilience agency. Link to the website: https://www.agriculture.gov.au/abares/research-topics/aboutmyregion/wa.
61. Ludwig, F., Milroy, S. & Asseng, S. Impacts of recent climate change on wheat production systems in western australia. *Climatic Change* **92**, 495–517 (2009).
62. Boruff, B., Biggs, E., Pauli, N., Callow, N. & Clifton, J. Changing water system vulnerability in western australia's wheatbelt region. *Applied Geography* **91**, 131–143 (2018).
63. Kirichenko, L., Radivilova, T. & Deineko, Z. Comparative analysis for estimating of the hurst exponent for stationary and nonstationary time series. *Information Technologies & Knowledge* **5**, 371–388 (2011).
64. Saji, N. H., Goswami, B. N., Vinayachandran, P. N. & Yamagata, T. A dipole mode in the tropical indian ocean. *Nature* **401**, 360–363 (1999).
65. Climate data online. Australian Government Bureau of Meteorology, Link to the website: http://www.bom.gov.au/climate/enso/history/ln-2010-12/IOD-what.shtml.
66. NSIDC. Daily sea ice extent (2020). Data can be downloaded from: https://psl.noaa.gov/repository/entry/show?entryid=f8d470f4-a072-4c1e-809e-d6116a393818.
67. NOAA. Daily nao index since january 1950 (2020). Data can be downloaded from: https://nsidc.org/data/G02135/versions/3.

Nested Monte Carlo simulation for molecular systems: Current status and future directions

Pradipta Bandyopadhyay*

School of Computational and Integrative Sciences

Jawaharlal Nehru University

New Delhi, INDIA 110067

[**Abstract:** Computer simulation techniques are predominantly used to study many-particle systems having complex interactions. Although the simulation techniques are well developed, one bottleneck for their application to larger systems is high computational cost. Along with the improvement of software and hardware, the improvement of simulation techniques plays an important role in pushing the limit of computer simulations. The Nested Monte Carlo (NMC) technique is a promising technique that increases the efficiency of the basic Monte Carlo (MC) simulation. To study a particular statistical mechanical ensemble, MC simulation generates a Markov chain of states having the limiting distribution the same as that of the ensemble under consideration. In the NMC method, the primary Markov chain is aided by an auxiliary Markov chain with a simpler energy function. The use of a simpler energy function can make the simulation faster. In this review, the development of this method from the first principle is described and the applications of the NMC techniques in molecular and biomolecular simulations are described. The review also discusses the technical challenges and possible ways to improve this promising method.]

*__Email__: praban07@gmail.com

1. Introduction

Statistical Mechanics is the bridge between mechanics (classical or quantum; in this manuscript, we will discuss only the classical part) and thermodynamics [1,2]. It can also be thought of as the theory for many-particle systems when all information about the system is not known. The theory for equilibrium statistical mechanics is well developed. However, for many of the problems, the calculation of the partition function is not possible due to complex interactions among the particles. Usually, for atomic and molecular systems numerical simulations are used to calculate the quantities of interest [3,4]. Molecular dynamics (MD) and Monte Carlo (MC) are the two most important computer simulation techniques to calculate different properties for a many-body system comprising atoms and molecules. MD solves Newton's equation to move in the phase space of 6N coordinates for the N particle system. On the other hand, MC aims to estimate the average values of quantities by generating states of the statistical mechanical ensemble, under consideration, using importance sampling. For the simulation of biomolecules such as proteins and nucleic acids, MD is usually the method of choice [5]. However, MC has its advantage being more flexible and can be more suitable than MD in some problems [6,7]. One of the primary issues for both MD and MC is the significant computational cost involved. There have been multiple attempts to reduce the computational cost while getting useful information from simulations. One attempt is to reduce the number of particles in the system by combining different particles into one particle, the so-called coarse-grained model [8], another is using a continuum model for solvation instead of explicit water molecules [9, 10]. A general way is to modify the simulation techniques using different tweaks and tricks [11-20]. In this review, a specific technique, Nested Monte Carlo (NMC), will be described that reduces the computational time [21]. For completeness, we will derive the working equations of a standard MC simulation before going to the details of the NMC techniques. Representative examples of the NMC techniques will be given. This review will end with possible directions to improve this technique.

2. Methodology

In the MC simulation, a distribution of first order Markov states (henceforth, Markov states) is generated with the condition that this distribution must follow the distribution for the statistical mechanical ensemble under consideration. A first order Markov chain (henceforth, Markov chain) depends only on the previous state and the total number of states are finite [3,4].

Any state in a Markov chain can be reached from any other state and the reverse is also true i.e. from any state all other states can be reached. Then the flux from state i to all other states can be given by

$$\sum_j \rho_i P_{ij} \tag{1}$$

The flux to i from all other states is given by $\sum_j \rho_j P_{ji}$ where ρ_i is the probability of the state i (e.g. in canonical ensemble $\rho_i \propto \exp(-\beta E_i)$, $\beta = \frac{1}{kT}$, k and T are Boltzmann constant and absolute temperature, respectively and E_i is the energy of the i-th state) and P_{ij} is the transition probability from state i to state j. At equilibrium, these two fluxes are the same and we can write

$$\sum_j \rho_i P_{ij} = \sum_j \rho_j P_{ji} \tag{2}$$

In MC simulations usually a more stringent condition is imposed where for every pair of states the condition given in eq. (2) is enforced.

$$\rho_i P_{ij} = \rho_j P_{ji} \tag{3}$$

In the canonical ensemble using this condition of microscopic reversibility, a distribution of states is generated having the limiting distribution as Boltzmann distribution. The recipe to ensure that the correct probability distribution is generated is given below.

The transition probability is given by

$$P_{ij} = \alpha_{ij} acc_{ij}$$

where α_{ij} is the probability of moving from state i to state j and acc_{ij} is the probability of acceptance of the state j from i. Equation (3) can be written as

$$\rho_i \alpha_{ij} acc_{ij} = \rho_j \alpha_{ji} acc_{ji} \tag{4}$$

In a standard MC simulation α is taken to be symmetric and hence eq. (4) becomes

$$\rho_i acc_{ij} = \rho_j acc_{ji} \tag{5}$$

$$\frac{acc_{ij}}{acc_{ji}} = \frac{\rho_j}{\rho_i}$$

As per Hastings-Metropolis criterion [3,4] the acceptance is considered as

$$acc_{ij} = \min\left(1, \frac{\rho_j}{\rho_i}\right) \tag{6}$$

For actual calculations in a metropolis sampling (eq. 6) Monte Carlo simulation, one needs to start from any state of a system and then change the state of the system (there is no restriction on how to change the state of a system; for instance, for a molecular system, one can change the coordinates of one or a few particles). Then eq. (6) is used to check the value of the acceptance. If $\rho_j > \rho_i$ then the right hand side of eq. (6) is one and the new state is accepted. If not, then the value of acceptance is a fraction i.e. the trial state has X percent chance of being accepted. To ensure that, a random number between 0 and 1 is generated uniformly and acceptance ratio is compared with that random number. If the acceptance ratio is more than or equal to the random number then the state is accepted; otherwise, the old state becomes the new state. If it is done for a large number of steps, this recipe ensures a correct Boltzmann distribution.

Although the basic idea of MC simulation is clear, it is often computationally too expensive and this prohibits the use of it in many interesting applications. There is a continuous effort to make the MC simulation faster. NMC is one of these techniques that holds considerable promise.

In the NMC method, two MC chains are generated; one with a cheaper potential energy function (auxiliary chain), and the other is the potential energy function of interest (primary chain). Here potential energy function (PEF) simply means the potential energy of different states of the system under consideration. The new state of the primary chain is generated by running a short MC simulation in the auxiliary chain. This can reduce the computational cost of the simulation significantly. This methodology is shown as a flowchart in figure 1 and schematically in Figure 2. The justification of this scheme is given below.

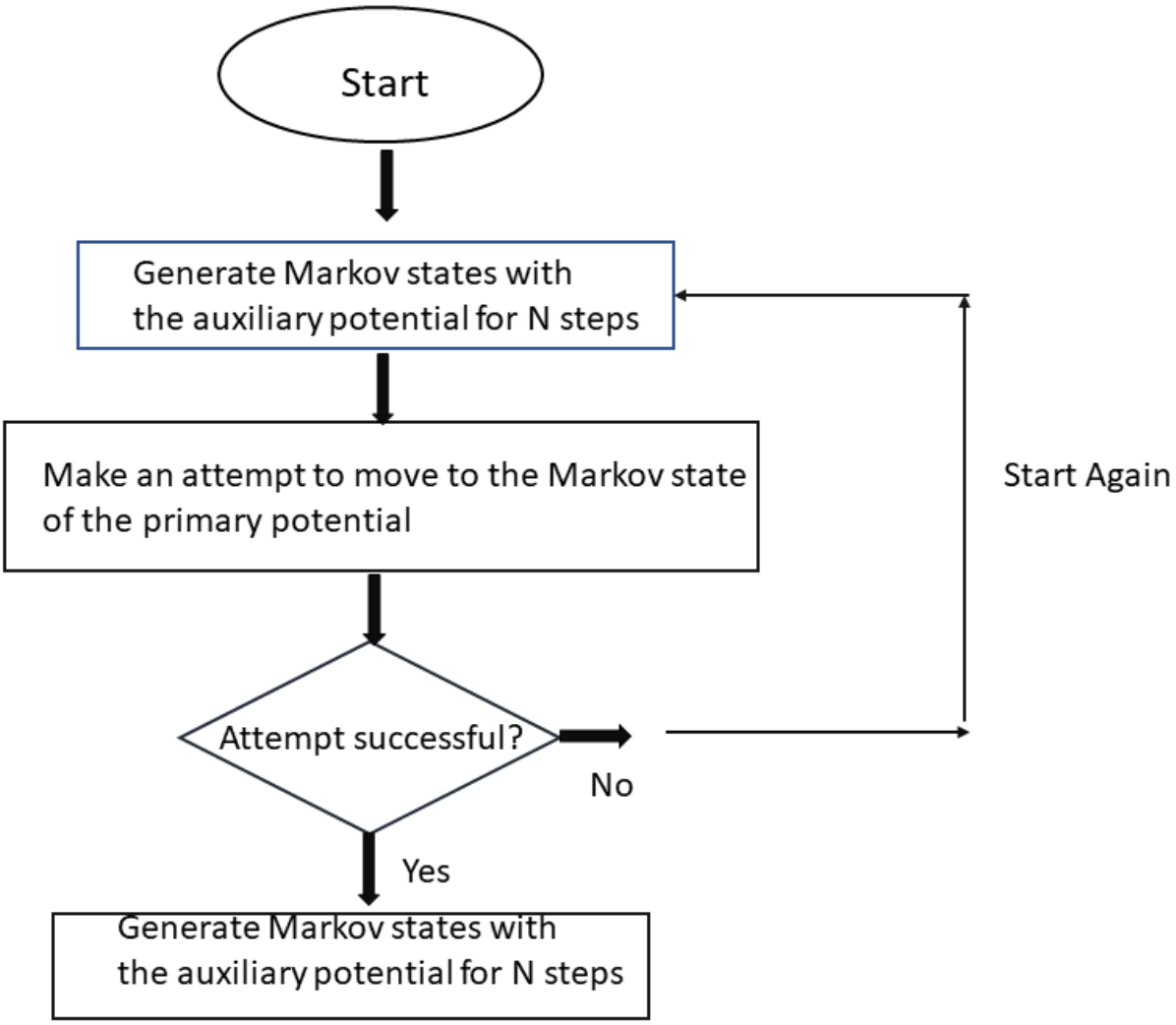


Figure 1: Flowchart for one cycle of the NMC method

In the NMC method, α_{ij} the probability of changing the state i to j is not taken as symmetric. Here the move to generate the new configuration is done by running a new MC Markov chain with a simpler PEF. The idea is to make the simulation faster as in this algorithm

most of the time will be spent in propagating the simpler Markov chain. The primary chain will be denoted as expensive (exp) and the auxiliary chain will be denoted as cheap (ch).

Hence, eq. (4) can be written with the new notation (also, i and j are replaced by x and y, respectively)

$$\rho_x^{\exp}\pi_{xy}^{\exp} = \rho_y^{\exp}\pi_{yx}^{\exp} \tag{7}$$

$$\rho_x^{\exp}\alpha_{xy}^{\exp}acc_{xy}^{\exp} = \rho_y^{\exp}\alpha_{yx}^{\exp}acc_{yx}^{\exp} \tag{8}$$

Now the probability of the move (α) is given by a short run using the cheap potential. For the Markov chain for the cheap potential one can write

$$\rho_x^{ch}\pi_{xy}^{ch} = \rho_y^{ch}\pi_{yx}^{ch} \tag{9}$$

Now for the expensive Markov chain the probability of generating the next state is given by

$$\alpha_{xy}^{\exp} = \pi_{x,1}^{ch}\pi_{1,2}^{ch}\ldots\pi_{y-1,y}^{ch}$$

and the same for the reverse move,

$$\alpha_{yx}^{\exp} = \pi_{y,y-1}^{ch}\pi_{y-1,y-2}^{ch}\ldots\pi_{1,x}^{ch}$$

$$\frac{\alpha_{xy}^{\exp}}{\alpha_{yx}^{\exp}} = \frac{\pi_{x,1}^{ch}\pi_{1,2}^{ch}\ldots\pi_{y-1,y}^{ch}}{\pi_{y,y-1}^{ch}\pi_{y-1,y-2}^{ch}\ldots\pi_{1,x}^{ch}}$$

$$= \frac{\rho_1^{ch}\rho_2^{ch}\ldots\rho_y^{ch}}{\rho_{y-1}^{ch}\ldots\rho_1^{ch}\rho_x^{ch}}$$

$$= \frac{\rho_y^{ch}}{\rho_x^{ch}} \tag{10}$$

The expression previous to equation (10) was obtained from the condition of detailed balance in the auxiliary chain. Now, the total acceptance can be written from equation (8) and (10)

$$\frac{acc_{xy}^{\exp}}{acc_{yx}^{\exp}} = \frac{\rho_y^{\exp}}{\rho_x^{\exp}} \frac{\rho_x^{ch}}{\rho_y^{ch}} = \frac{\exp(-\beta E_y^{\exp})}{\exp(-\beta E_x^{\exp})} \frac{\exp(-\beta E_x^{ch})}{\exp(-\beta E_y^{ch})} \tag{11}$$

$$acc_{xy}^{\exp} = \min(1, \exp(-\beta(\Delta E^{\exp})) \times \exp(\beta(\Delta E^{ch}))) \tag{12}$$

Where $\Delta E = E_y - E_x$

However, there are challenges in using eq. (12). This acceptance condition says that the difference between energies in the x and y state should be similar for expensive and cheap chains to get good acceptance.

In practical terms, it is necessary to think about an auxiliary potential energy function so that the rate of acceptance between cheap and expensive Markov Chains will be reasonable. If the acceptance is too low then this algorithm will not work. More on this issue is discussed in the next section.

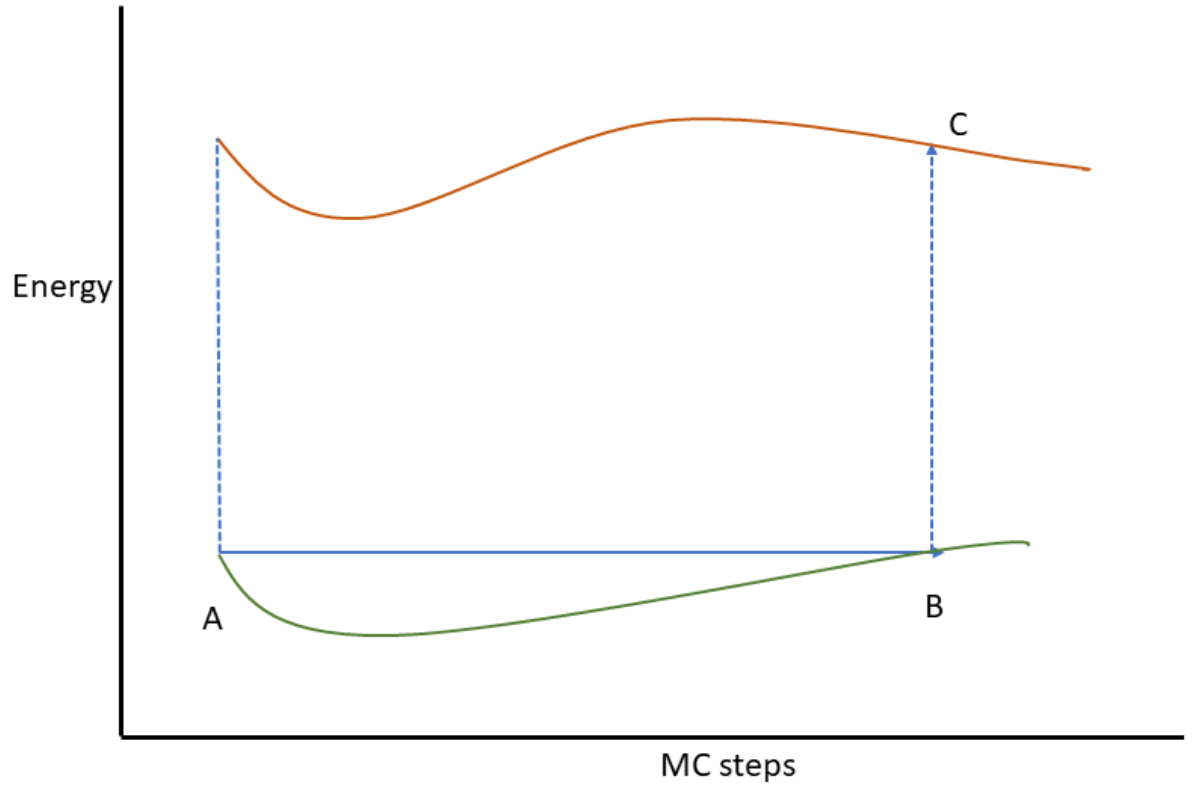


Figure 2: A schematic representation of the nested Monte Carlo (NMC) method. The top curve is the primary chain, while the bottom is the auxiliary chain. The propagation of the system is done in the auxiliary chain (A to B) and a new state C on the primary chain is obtained.

3. Applications

One of the first applications of the NMC method was done by Muller and Warshel to calculate the free energy barriers for chemical reactions in solution [21]. In their case the auxiliary potential was taken from the empirical valence bond (EVB) method and the primary potential from *ab initio* (Hartree-Fock) method. Gelb has used this method with Lennard-Jones (LJ) potential as a test case,

where the full LJ potential is the primary potential, while cut and shifted potential is the auxiliary potential [22]. Luo has used it for protein systems, where the primary and auxiliary potentials are taken as all atom and coarse-grained potentials, respectively [23]. Bandyopadhyay has used NMC in several works [24-26], with applications on water clusters, where the primary potential is quantum mechanical and the auxiliary potential is classical. In one work Bandyopadhyay used quantum mechanical/molecular mechanical (QM/MM) potential for the primary chain and a molecular mechanical (MM) potential as the auxiliary chain [24]. Other groups such as Iftmie et al., Matusek et al. also used to connect MM and QM methods using this algorithm [27,28]. Calvo has used it by combining polarizable and non-polarizable potentials [29]. It is also used in molecular dynamics (MD); for instance, The equilibrium thermodynamics of dense fluid nitrogen was investigated using this method [30].

Although the range of applications is quite impressive there are several drawbacks of this method. To understand it let us define two efficiency metrics.

$$speed\,up = \frac{number\ of\ \ steps * cost\_expensive}{\left(\frac{number\ of\ \ steps}{jump\ \ frequency}\right) * cost\_expensive + number\ of\ \ steps * cost\_cheap}$$

$$jump\_efficiency = \frac{Accepted\ \ jumps}{jumps\ attempted}$$

Here, cost_expensive and cost_cheap indicates the computational cost associated with the one step of the MC run with expensive and cheap potential energy functitons. Jump frequency is the ratio of the total number of steps in the auxiliary chain and the number of jump attempts made to move from the auxiliary chain to the primary chain and accepted jumps are the number of successful movements between auxiliary and primary chains.

We want the computational speedup to be maximum. However, for that the jump frequency should be low i.e. the jump between primary and auxiliary potentials should be attempted less often.

However, the increase of speedup and jump efficiency is inversely related. If the jumps are attempted too often there is more chance of acceptance but this increases the computational cost.

There have been attempts to optimize NMC; Coe et al. have optimized thermodynamic variables like T and P to get the maximum overlap between the primary and auxiliary potential functions [31]. Bandyopadhyay scaled the inverse temperature, β, used to propagate the auxiliary chain to get higher acceptance rate [32]. In a more recent work, Srivastava et al. have optimized the two distributions, primary and auxiliary, by minimizing the Kullback-Liebler (KL) divergence between them [33].

3.1 Remaining Issues

The efficiency of this algorithm can be optimized even more. Already there are attempts to do incremental coarse-graining. In the work from the group of Zuckerman, they have changed the energy functiton incrementally with the small change between two consecutive potential energy functions [34]. The group of Juan De Pablo also developed an algorithm where a layered nested MC chain is used [35]. In both of these methods, as the two-layered NMC is replaced by a multi-layer NMC, the difference of potential functions reduces between two consecutive layers giving rise to higher acceptance. However, the evaluation of a larger number of potential energy functions may increase the computational cost of the algorithm. The NMC method has got special impetus in recent years due to the advancement of machine learning (ML) techniques for molecular systems. One important application is this area is to develop faster potential energy functions using ML techniques. However, if the ML based PEF can be thought as the auxiliary potential then they do not need to be super accurate to the original PEF. Of course, the ML based PEF should have reasonable similarity with the primary PEF so that the acceptance between these PEFs become large enough for getting meaningful statistical averages. Jadrich et al. have developed an NMC scheme where the primary potential function is based on Density Functional Theory (DFT) while the auxiliary chain comes from a ML potential [36]. They have used this method for investigating high-explosive materials.

3.2 New Ideas that can be tried

Srivastava et al. in reference 33 have studied symmetric and asymmetric ionic systems where the cheap potential function was improved using the KL divergence between primary and auxiliary

distributions. A more general approach can be thought of in the following way. The simulation starts with a reasonable auxiliary potential function and its parameters are periodically updated during the simulation (on-the-fly). However, it can be envisaged that initially, the jump frequency has to be large so that there will be enough data to parametrize the auxiliary potential. It can be used for a small number of degrees of freedom. Mathematically the minimization of the KL divergence (D_KL) between the two distributions is shown below. $p(x_i)$ is the (reduced) distribution function coming from the primary chain, while $q(x_i,\lambda)$ is that coming from the auxiliary chain. λ is the set of parameters that can be optimized to get a lower value of the KL divergence. x is a subset of all coordinates of the system and i denotes the states of the system.

$$\min\left(D_{KL}(p \,||\, q) = \sum_{i=1}^{N} p(x_i) log \frac{(p(x_i))}{(q(x_i;\lambda))} \right)$$

The algorithm is described below.

1. Initially run N number of short runs (where the jump frequency is one) having different λ values and calculate D_KL for each case. Estimate the value of λ for which D_KL is the minimum. However, as these runs are short the estimate of D_KL is only approximate.
2. In the second part of the algorithm run longer simulations with M (M<N) λ parameters, where these λ values are close to the one that gave minimum D_KL in the previous cycle, and again estimate the value of λ for the minimum value of D_KL. This λ can be a good starting point for the auxiliary potential. A schematic representation of this method is given in Figure 3.
3. The next step is to optimize the jump frequency. Here, the λ is kept fixed but the jump frequency is varied. An optimum value of the jump frequency is to be chosen while keeping the KL divergence low. This idea can be used to improve the acceptance rate in NMC simulations.

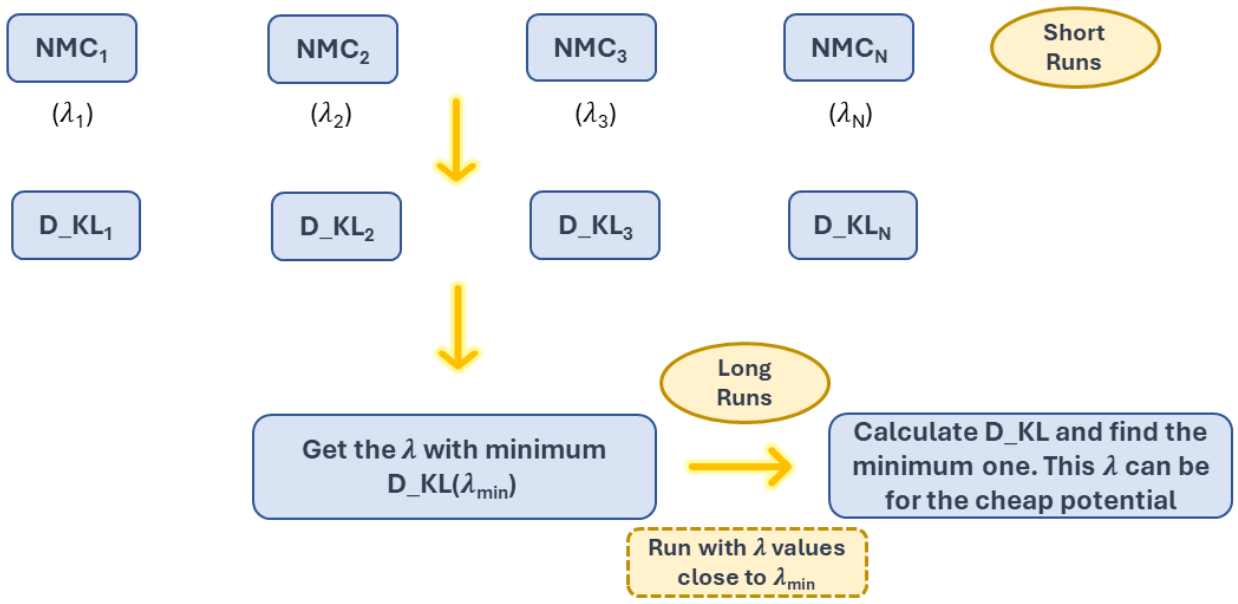


Figure 3: A proposed scheme to obtain a good auxiliary potential using Kullback-Liebler (KL) divergence. In the first cycle, short NMC runs with different auxiliary potentials are performed (lambda is a set of parameters controlling the auxiliary potential). In the second cycle, lambda values are chosen around the value that gave minimum KL divergence in the previous cycle. Longer NMC runs can be performed to get better λ and a reasonable auxiliary potential.

4. Conclusion

In this review article, the details of the Nested Monte Carlo technique have been described. This technique has the potential to improve the efficiency of molecular simulation and has already been used in areas like chemical reactions in solutions, molecular clusters, protein structure and dynamics. It has also been used to run quantum mechanical and quantum mechanical/molecular mechanical (QM/MM) simulation using MM and machine learning based potentials as importance functions. The application of this method and technical challenges are illustrated with possible avenues for improvement.

5. Acknowledgement

The author would like to thank Dr. Rakesh Srivastava of VIT Bhopal for illuminating discussions and both Pooja and Alakananda Banerjee for help in making the figures.

References

1. L.D.Landau and E. M. Liftshitz, Statistical Physics, 1970, Pergamon Press.

2. M. Kardar, Statistical Physics of Particles, 2007, Cambridge University Press, .

3. M. P. Allen and D. J. Tildesley, Computer simulation of liquids: Second edition, 2017, Oxford University Press.

4. D. Frenkel and B. Smit, Understanding molecular simulation: From algorithms to applications, 2001, Academic Press; 2nd edition.

5. A. Hospital, J. R. Goñi, M. Orozco, J. L. Gelpí, Advances and Applications in Bioinformatics and Chemistry, 2005, 8, 37.

6. K. Binder, D. W. Heermann, Monte Carlo Simulation in Statistical Physics: An Introduction, 2019, Springer.

7. D. Zuckerman, Annual Rev. Biophysics, 2011, 40, 41.

8. J. Jin, J. P. Alexander, A. E. P. Durumeric, T. D. Loose, G. A. Voth, Journal of Chemical Theory and Computation 2022, 18, 5759

9. W. C. Still, A. Tempczyk, R.C.Hawley, T. Hendrickson, J. Am. Chem. Soc. 1990, 112, 6127.

10. L. Xiao, C. Wang, R. Luo, Journal of Theoretical and Computational Chemistry, 2014, 13, 1430001.

11. Y. Sugita, Y. Okamoto, Chem. Phys. Lett., 1999, 314, 141.

12 F. Wang, D. P. Landau, Phys. Rev. Lett., 2001, 86, 2050.

13. P. Singh, S. Sarkar, P. Bandyopadhyay, Chem. Phys. Lett, 2011, 514,357.

14. A. Barducci, M. Bonomi and M. Parrinello, WIREs Comput Mol, 2011, 1, 826.

15. F. Yasar, A.J. Ray and U.H.E. Hansmann, Phys. Rev. E, 2022, 106, 015302.

16. C. Liu, E. Brini, A. Perez, K. A. Dill, Journal of Chemical Theory and Computation, 2020, 16, 6377.

17. J. Wang, P.R. Arantes, A. Bhattarai A, 2021. WIREs Comput Mol Sci., 11, e1521.

18. W. Jiang, B. Roux, J. Chem. Theory Comput., 2010, 6, 2559.

19. H. Sidky, W. Chen, A.L. Ferguson, Molecular Physics, 2020, 118(5) doi: 10.1080/00268976.2020.1737742.

20. Y, Wang, J. M. L. Ribeiro, P. Tiwary, Current Opinion in Structural Biology, 2020, 61, 139.

21. R. P. Muller, A. Warshel, Journal of Physical Chemistry, 1995, 99, 17516.

22 L. D. Gelb, Journal of Chemical Physics, 2003, 118, 7747.

23. T. Z. Lwin, R. Luo, Journal of Chemical Physics, 2005, 123, 194904.

24 P. Bandyopadhyay, Journal of Chemical Physics, 2005, 122, 091102.

25 P. Bandyopadhyay, Theor Chem Acc, 2008, 120, 307.

26 P. Bandyopadhyay, Journal of Chemical Physics, 2008, 128, 134103.

27. R. Iftimie, J. Schofield, Journal of Chemical Physics, 2001, 114, 6763.

28. D. R. Matusek, S. Osborne, A. St-Amant, J. Chem. Phys., 2008, 128, 154110.

29. F. Calvo, Int J Quantum Chem, 2010, 110, 2347.

30. J. D. Coe, T. D. Sewell and M. S. Shaw, Journal of Chemical Physics, 2009, 130, 164104.

31. J. D. Coe, T. D. Sewell and M. S. Shaw, J. Chem. Phys., 2009, 131, 074105.

32. P. Bandyopadhyay, Chem Phys Lett, 2013, 556, 341.

33. R. Srivastava, R., P. Bandyopadhyay, J Chem Sci, 2023, 135, 51.

34. N. E. Jackson, M. A. Webb and J. J. de Pablo, J Chem Phys, 2018, 149, 072326.

35. E. Lyman, F. M. Ytreberg and D. M. Zuckerman, Phys Rev Lett, 2006, 96, 028105.

36. R. B. Jadrich and J. A. Leiding, J. Phys. Chem, 2020, 2020, 5497.

Complexity in the distribution of the dynamics of foraging ants in an ecosystem

R.K. Brojen Singh*

School of Computational & Integrative Sciences,

Jawaharlal Nehru University, New Delhi-110067, India.

Abstract: The dynamics of foraging ants in a sub-populated colony in a certain ecosystem is a very complex phenomenon which involve a continuous interaction of the ants subjected to two identical food sources. This complicated interaction of the ants allows the ants to follow in different trajectories leading to one of the two food sources. We study a simple model of continuously interacting two groups of ants in search of food leading to forage the two identical food sources. We then constructed a Master equation of the model and followed van Kampen's system size expansion to approximate the model equation to a linear Fokker-Planck equation from which we solved probability distribution of the single trajectory of the foraging ants of a certain type. The asymptotic properties of the probability distribution indicates that at sufficiently long time, the probability distribution of the foraging ants becomes constant showing a possible tendency to go to a fixed size of the foraging ants.

Keywords: Foraging ants; Stochastic process; Master equation; Kramers-Moyal expansion; Kampen's system size expansion.

*Electronic address: brojen@jnu.ac.in

I. INTRODUCTION

Simple interaction may result in complicated dynamics and processes. Mathematically, interaction effects comes into existence when the effect of one variable depends on other interacting variable. [1]. This interaction process may exhibit interesting dynamical phenomena, such as, multistability including bistability, oscillation, chaotic dynamics and many other phenomena [1]. The foraging ants of a certain colony of ants to two identical sources of food is one example of simple model which is experimentally well studied [2, 3]. They showed that this foraging ants are distributed with different sizes of the populations along the two food sources. They showed that after sufficiently long time, the ants switch to forage to the other food source [3, 4]. However, these behaviours of the foraging ants are not fully studied.

There have been simple models to address the possibility of bistability [9] which was observed experimentally [3, 4]. They showed that even though the deterministic model could not able to exhibit this bistability phenomena, the stochastic model simulation results clearly showed it. Generally, the population of ants in a certain colony is finite and small as compared to the whole population in the ecosystem. In such situation, the dynamics of these ants are driven by fluctuations (noise) arising from random interaction among the ants in the colony (intrinsic fluctuations) and random interaction of the ants with the surrounding environment (extrinsic fluctuations) [5, 6]. These fluctuations cannot be neglected but play important roles, such as, driving the system to nonequilibrium state [7], in important decision making by them [6], in maintaining collective behaviour [8] and many others. In some situations, these fluctuations could be responsible for various dynamical properties, namely, bistability phenomenon [9]. However, the behaviour of the distribution function of the foraging ants system was not addressed properly and fully.

Unlike deterministic method, stochastic approach to understand the behaviour of the model systems of interacting molecules is quite important as the approach systematically takes into account the effect of fluctuations and how it drives the system to various states [10, 11]. Stochastic approach allows to construct a model equation known as Master equation of any model system consisting of randomly interacting molecules which undergo a certain number of reaction channels [10, 11]. In general, this multi-variable Master equation cannot

be solved analytically except for simple systems. However, numerically one can solve such system consisting of a number of reaction channels representing interaction of the molecular species in the system using stochastic simulation algorithm or Gillespie algorithm [12]. Analytically, there have been approximate methods to solve Master equation [10, 11]. The van Kampen's system size expansion method is one of the approximate methods to solve Master equation by modeling the transition probability in the Master equation to be dependent on the size of the jump [10]. This approximation allows to express intensive variables to extensive one which can be expanded in terms of system's size. This method was shown to be quite accurate in comparison to other methods [13]. In this work, we use this van Kampen's system size expansion method to solve and addressed the behaviour of the probability distribution function of the foraging ants system, it's asymptotic behaviour and found to be quite in agreement with the experimentally found results in [3, 4].

We studied a simple model of foraging ants by Biancalani et. al. [9] and solved the equivalent Master equation using van Kampen's system size expansion approach. The solved probability distribution function is analyzed and found to be dependent on time and system size which is closely in agreement with the experimentally observed results [3, 4]. The work is organized as in the following, first we introduced the van Kanpen's system size expansion method to calculate probability distribution function, next results on the foraging ants are presented, and lastly conclusion is given based on the results we obtained.

II. PRELIMINARY SET UP AND NOTATION

Consider a random physical variable x which have Markovian character, then the dynamics of its probability distribution $P(x,t)$ satisfy the following Master equation [10],

$$\frac{\partial P(x,t|x',t')}{\partial t} = \int [W(x|x')P(x',t) - W(x'|x)P(x,t)] \tag{1}$$

This Master equation (1) is the jump approximation of the Chapman-Kolmogorov equation in differential form [11]. Here, $W(x|x')$ is the transition per unit time from the state x' to the state x. Now consider $y = x - x'$ for the first term and $y = x' - x$ for the second term, then one can write transition rate as $t(y,x) = W(x+y|x)$. Now the equation (1) can be

expanded by the following Kramers-Moyal expansion [14, 15],

$$\frac{\partial P(x,t|x',t')}{\partial t} = \sum_{n=1}^{\infty} \frac{1}{n!} \left[-\frac{\partial}{\partial y}\right]^n [\alpha_n(x) P(x,t|x',t')] \tag{2}$$

where, the parameter $\alpha_n(x)$ is given by,

$$\alpha_n(x) = \int dx' \, [x'-x]^n \, W(x'|x) \tag{3}$$

The van Kampen's power series expansion [16] allows to define extensive variable x (such as number of molecules etc) and intensive variable u (concentration of molecules $u = \frac{x}{\Omega}$ which is independent of system's size Ω) to approximate $W(x'|x)$ as follows. The transition rate $W(x'|x)$ from state x to x' is directly proportional to the jump size $\Delta x = x' - x$. If we scale the extensive variable x by extensive variable $\frac{x}{\Omega}$, then $W(x'|x)$ can be scaled as follows,

$$W(x'|x) \propto W(x;\Delta x) = \Omega\phi\left(\frac{x}{\Omega};\Delta x\right) \tag{4}$$

In such situation, ϕ is independent of Ω. Now, scaling $x \to \frac{x}{\Omega}$ and $x' \to \frac{x'}{\Omega}$ to equation (4), and then putting to the equation (3), we have,

$$\alpha_n(x) = \Omega \int d[\Delta x][\Delta x]^n \phi\left(\frac{x}{\Omega};\Delta x\right) = \Omega\tilde{\alpha}_n\left(\frac{x}{\Omega}\right) \tag{5}$$

Now, this $\tilde{\alpha}_n$ is independent of Ω. From this analysis, Kampen introduced the idea of possibility of expressing the variable x in terms of normalized variable $u = \frac{x}{\sqrt{\Omega}}$ [16], such that,

$$x = \Omega\phi(t) + u\sqrt{\Omega} \tag{6}$$

where, the function $\phi(t)$ is to be determined. Now, using this variable transformation (6) and putting in the equation (2) and after simplification, we have,

$$\begin{aligned} \frac{\partial P(u,t|u',t')}{\partial t} &- \sqrt{\Omega}\phi'(t)\frac{\partial P(u,t|u',t')}{\partial u} \\ &= \sqrt{\Omega}\left[-\frac{\partial}{\partial u}\right]\tilde{\alpha}_1\left[\phi(t)+\frac{u}{\sqrt{\Omega}}\right]P(u,t|u',t') \\ &\quad + \sum_{n=2}^{\infty}\frac{\Omega^{1-\frac{n}{2}}}{n!}\left[-\frac{\partial}{\partial u}\right]^n\left[\tilde{\alpha}_n\left\{\phi(t)+\frac{u}{\sqrt{\Omega}}\right\}P(u,t|u',t')\right] \end{aligned} \tag{7}$$

In the large Ω limit: $\lim_{\Omega\to\infty}\frac{x}{\Omega} = \lim_{\Omega\to\infty}\left[\phi(t)+\frac{u}{\sqrt{\Omega}}\right] = \phi(t) = finite$ (using equation (6)), which is known as the *thermodynamics limit* recovering deterministic or macroscopic behaviour [17]. Now, equating the co-efficient of $\Omega^{1/2}$ from both sides of the above equation

(7), and taking large Ω limit: $\Omega \to \infty$ (macroscopic limit), we have,

$$\begin{aligned} -\phi'(t)\frac{\partial P(u,t|u',t')}{\partial u} &= \lim_{\Omega\to\infty}\left[-\frac{\partial}{\partial u}\right]\tilde{\alpha}_1\left[\phi(t)+\frac{u}{\sqrt{\Omega}}\right]P(u,t|u',t') \\ &= -\tilde{\alpha}_1[\phi(t)]\frac{\partial P(u,t|u',t')}{\partial u} \end{aligned} \tag{8}$$

From which one can immediately retrieve macroscopic dynamics given below,

$$\frac{d\phi(t)}{dt} = \tilde{\alpha}_1\{\phi(t)\} = \frac{1}{\Omega}\alpha_1\{\phi(t)\} \tag{9}$$

Now, expanding $\tilde{\alpha}_n\left[\phi(t)+\frac{u}{\sqrt{\Omega}}\right]$ in powers of $\Omega^{-\frac{1}{2}}$ and rearranging the terms, we have,

$$\frac{\partial P(u,t|u',t')}{\partial t} = \sum_{m=2}^{\infty}\frac{\Omega^{-\frac{m-2}{2}}}{m!}\sum_{n=1}^{m}\frac{m!}{n!(m-n)!}\tilde{\alpha}_n^{m-n}\left[-\frac{\partial}{\partial u}\right]^n u^{m-n}P(u,t|u',t') \tag{10}$$

Now, if we take large Ω limit i.e. $\Omega \to \infty$, then, only the two terms for $m = 2$ which are independent of Ω will survive and all the rest terms will be vanished. The reduced equation at this condition becomes,

$$\frac{\partial P(u,t|u',t')}{\partial t} = -\tilde{\alpha}_1'\{\phi(t)\}\frac{\partial}{\partial u}[uP(u,t|u',t')] + \frac{1}{2}\tilde{\alpha}_2\{\phi(t)\}\frac{\partial^2}{\partial u^2}P(u,t|u',t') \tag{11}$$

Next, in order to obtain the expression for $\tilde{\alpha}_1\{\phi(t)\}$, initial distribution at $t = 0$ is taken as $P(u,0) = \delta(u-u_0)$. Then define a parameter s as,

$$s(t) = ln\left[\frac{\tilde{\alpha}_1\{\phi(0)\}}{\tilde{\alpha}_1\{\phi(t)\}}\right] \tag{12}$$

so that $s(0) = 0$. Then, differentiating (12) with respect to t and using equation (9), we have,

$$\begin{aligned} \frac{ds(t)}{dt} &= \frac{d}{dt}\left[ln\tilde{\alpha}_1\{\phi(0)\} - ln\tilde{\alpha}_1\{\phi(t)\}\right] \\ &= -\frac{1}{\tilde{\alpha}_1\{\phi(t)\}}\frac{d}{d\phi}\tilde{\alpha}_1\{\phi(t)\}\frac{d\phi(t)}{dt} \\ &= -\frac{1}{\tilde{\alpha}_1\{\phi(t)\}}\frac{d}{d\phi}\tilde{\alpha}_1\{\phi(t)\}\times.\tilde{\alpha}_1\{\phi(t)\} \\ &= -\frac{d}{d\phi}\tilde{\alpha}_1\{\phi(t)\} \end{aligned} \tag{13}$$

Now, coordinate transformation of equation (11) with respect to s is done $P(u,t|u',t') \to P(u,s|u',s')$. Then, if transformations of $u = ye^{-s}$, and $P(u,s|u',s') = e^sH(y,s|y',s')$ are done, the equation (11) reduces to the following heat equation,

$$\frac{\partial H(y,s|y',s')}{\partial s} = -\frac{\tilde{\alpha}_2\{\alpha(t)\}}{2\tilde{\alpha}_1\{\phi_1(t)\}}e^{2s}\frac{\partial^2 H(y,s|y',s')}{\partial y^2} \tag{14}$$

This equation is analogous to heat equation which can be solved as an initial valued problem. Using Fourier transform technique, we can solve the equation (14) and then putting back the transformed function given above, we get,

$$P(u,s|u',s') = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-u'e^{-s})^2}{2\sigma^2}} \tag{15}$$

where,

$$\sigma^2 = e^{-2s}\int_{t'}^{t} \tilde{\alpha}_2\{\phi(\tau)\}e^{2s(\tau)}d\tau \tag{16}$$

Now, the initial distribution is given by, $P(x,0) = \delta(x-x') = \delta(x-\Omega u')$ and taking $\phi(0) = u'$ and using coordinate transformation relations, the distribution function (17) becomes,

$$P(x,t|x',t') = \frac{1}{\sqrt{2\pi\sigma^2\Omega}} e^{-\frac{[x-\Omega\phi(t)]^2}{2\sigma^2}} \tag{17}$$

This probability distribution function obtained using van Kampen's system's size expansion [16] is the distribution function of the single trajectory of the particle system defined by random variable x starting from an initial state (x',t'). The parameters σ^2 and $\phi(t)$ can be obtained by solving the equations (16) and (9) through the equation (12) for the system.

III. RESULTS

The model considers a colony of ants of finite population N subjected to two identical food sources [9]. The ants who choose food source 1 for collecting food is named as x_1 and the ants who collect food from the food source 2 is denoted by x_2. The ants gathering food from one food source can select ants of it's own type from the population of ants in the colony. Further, ants of the types x_1 and x_2 can interact among themselves and can switch from one type to another depending on the availability of the food at the two food sources. The model interaction of the foraging ants is given by the following reactions [9],

$$\begin{aligned} x_1 + x_2 &\xrightarrow{r} 2x_1 \qquad x_1 + x_2 \xrightarrow{r} 2x_2 \\ x_1 &\xrightarrow{k} x_2 \qquad\qquad x_2 \xrightarrow{k} x_1 \end{aligned} \tag{18}$$

where, r and k are the classical rates of the respective interaction. Also $x_1 + x_2 = N$, where, N is the total population of the ants in the colony. If n is the population state

of the individual type x_1, then the population state of individual x_2 is $N-n$. Now, the dynamics of the configurational probability $P(n,t|n',t')$ of the set of chemical reactions (18) representing the model interaction at any instant of time t starting from an initial state n_0 at time t_0 is given by the stochastic Master equation. To solve this Master equation for the probability distribution $P(n,t|n',t')$, we followed the van Kampen's system size expansion method [16] as explained in the previous section. For this, we define $\frac{n}{N}$ as the intensive variable. Following Kramers-Moyal expansion [14, 15] as given in the previous section and taking the limit $N \to \infty$, the Master equation reduces to linear order Fokker-Planck equation as given by,

$$\begin{aligned}\frac{\partial P(n,t|n',t')}{\partial t} &= \int \left[W(n|n')P(n',t|n,t') - W(n'|n)P(n,t|n',t')\right] dn'. \\ &= \sum_{i=1}^{\infty} \frac{1}{i!}\left[-\frac{\partial}{\partial n}\right]^i \alpha_i(n) P(n,t|n',t') \\ &= -\tilde{\alpha}_1'[\phi(t)]\frac{\partial}{\partial u}[uP(u,t|u',t')] + \frac{1}{2}\tilde{\alpha}_2[\phi(t)]\frac{\partial^2}{\partial u^2}P(u,t|u',t') \qquad (19)\end{aligned}$$

The first term is the drift term with drift coefficient $\tilde{\alpha}_1'[\phi(t)]u$, whereas, the second term denotes diffusive term with diffusive coefficient $\frac{1}{2}\tilde{\alpha}_2[\phi(t)]$ [18]. As shown in the previous section, the solution of this linear Fokker-Planck equation using van Kampen's system size expansion is given by,

$$P(n,t|n',t') = \frac{1}{\sqrt{2\pi\sigma^2 N}} e^{-\frac{[n-N\phi(t)]^2}{2\sigma^2}} \qquad (20)$$

Now, in order to calculate ϕ and σ^2 in the equation (20), we have to calculate the transition probability $W(x|x')$ for the model reaction given in (18) and is given by,

$$\begin{aligned}W(n|n') &= [N-(n-1)]\left[\frac{r}{N}(n-1)+k\right]\delta_{n,n'-1} \\ &\quad +(n+1)\left[\frac{r}{N}\{N-(n+1)\}+k\right]\delta_{n,n'+1} \qquad (21)\end{aligned}$$

Now, let us take $\Delta n = n' - n$. Then, following van Kampen's system size expansion as discussed in the previous section, it can be approximated $W(n|n') \approx W(n';\Delta n)$. From the above equation (21), we can easily calculate $W(n'|n)$ from the delta function in the equation as given below,

$$\begin{aligned}W(n';\Delta n) &= [N-(n-1)]\left[\frac{r}{N}(n-1)+k\right]\delta_{\Delta n,-1} \\ &\quad +(n+1)\left[\frac{r}{N}\{N-(n+1)\}+k\right]\delta_{\Delta n,1} \qquad (22)\end{aligned}$$

From this approximated transition probability given by equation (22), the parameter $\tilde{\alpha}_1(n)$ can be calculated from the first moment given by equation (5) putting $n = 1$ given by,

$$\begin{aligned}\tilde{\alpha}_1 &= r(n-1) + kN - \frac{r}{N}(n-1)^2 - k(n-1) - r(n+1) \\ &+ r\frac{(n+1)^2}{N} - k(n+1)\end{aligned} \tag{23}$$

Taking large $N-$limit, rearranging the terms and taking $n \to \phi$, the parameter $\tilde{\alpha}_1[\phi(t)]$ can be expressed in terms of $\phi(t)$. Now, using equation (9), one can have the following linear differential equation in $\phi(t)$,

$$\frac{d\phi(t)}{dt} + 2k\phi(t) \approx kN - 2r \tag{24}$$

This equation can be easily solved using integrating factor method, where, integrating factor is given by, $IF = e^{\int 2kdt} = e^{2kt}$. Then, multiplying the equation (24) by this IF, and rearranging the term, we get, $\frac{d}{dt}\left[\phi(t)e^{2kt}\right] = kN - 2r$ Now, integrating this equation, we get the solution for $\phi(t)$. The integration constant can be calculated from the initial valued problem where, at $t = 0$, $\phi(0) \to \phi_0$, The solution $\phi(t)$ is given by,

$$\phi(t) = \phi_0 e^{-2kt} + (kN - 2r)\left[1 - e^{-2kt}\right] \tag{25}$$

In order to calculate σ^2 in equation (20) using equation (16), we have to calculate s and $\tilde{\alpha}_2[\phi(t)]$ from the equations (12) and (5) respectively. From equation (12), substituting the expression for $\phi(t)$ and after simplification of the expression, we have,

$$\begin{aligned}e^s &= \frac{a}{b - ce^{-2kt}} \\ a &= \frac{N}{2} - \left[\phi_0 + \frac{r}{k}\right] \\ b &= \frac{N}{2} + k\left[N - \frac{2r}{k}\right] - \frac{r}{k} \\ c &= \phi_0 + kN - 2r\end{aligned} \tag{26}$$

The parameter $\tilde{\alpha}_2$ is calculated using the equation (5) by using the expression for transition probability $W(n';\Delta n)$ (22), $\tilde{\alpha}_2[\phi(t)] = \int d[\Delta n][\Delta n]^2 W(n';\Delta n)$ with $n' \to n$ through $\Delta n = \pm 1$. Then, we take large N limit and the expression for $\tilde{\alpha}_2[\phi(t)]$ by taking $n \to \phi(t)$ is given by,

$$\tilde{\alpha}_2[\phi(t)] \approx 2r\phi(t) + k(N+2) \tag{27}$$

Now, using the equations (26) and (12), the parameter σ^2 is calculated using the equation (16). The expression for σ^2 is given by,

$$\begin{aligned} \sigma^2 &= 2re^{-2s}I_1 + k(N+2)e^{-2s}I_2 \\ I_1 &= \int_0^t \phi(t)e^{2s}dt \\ I_2 &= \int_0^t e^{2s}dt \end{aligned} \tag{28}$$

The integral I_2 can be calculated using equation (26) and after rearranging the terms, we have,

$$\begin{aligned} I_2 &= \int_0^t \left[\frac{a}{b-ce^{-ekt}}\right]^2 dt \\ &= \frac{1}{2kb^2}\left[\frac{c^2}{b}\left(\frac{1}{c-b} - \frac{1}{ce^{2kt}-b}\right) - ln\left(\frac{b-c}{be^{2kt}-c}\right)\right] \end{aligned} \tag{29}$$

The integral I_1 can be calculated using the expression for ϕ_t in equation (25), after rearranging the terms, we have,

$$\begin{aligned} I_1 &= (\phi_0 - kN + 2r)\int_0^t e^{-2kt}\left[\frac{a}{b-ce^{-2kt}}\right]^2 dt + (kN-2r)I_2 \\ &= \frac{ra^2}{kc}\left(\frac{b-ce^{-2kt}}{a}\right)^2\left[\frac{1}{ce^{-2kt}-b} - \frac{1}{c-b}\right] \\ &\quad + \frac{1}{2kb^2}\left[(kN-2r)2r\right]\left(\frac{b-ce^{-2kt}}{a}\right)^2 \\ &\quad \times \left[\frac{c^2}{b}\left(\frac{1}{c-b} - \frac{1}{ce^{2kt}-b}\right) - ln\left(\frac{b-c}{be^{2kt}-c}\right)\right] \end{aligned} \tag{30}$$

Now, the expression for σ^2 can be obtained by substituting the expressions for I_2 and I_1 from the equations (29) and (30) respectively to the equation (16). The expression for σ^2 is given by,

$$\begin{aligned} \sigma^2(t) &= \frac{ra^2}{kc}\left(\frac{b-ce^{-2kt}}{a}\right)^2\left[\frac{1}{ce^{-2kt}-b} - \frac{1}{c-b}\right] \\ &\quad + \frac{1}{2kb^2}\left[(kN-2r)2r + k(N+2)\right]\left(\frac{b-ce^{-2kt}}{a}\right)^2 \\ &\quad \times \left[\frac{c^2}{b}\left(\frac{1}{c-b} - \frac{1}{ce^{2kt}-b}\right) - ln\left(\frac{b-c}{be^{2kt}-c}\right)\right] \end{aligned} \tag{31}$$

Now, the probability distribution of the population of ants of type $n = x_1$ can be obtained from the equation (20) with the expression (31). It is given by,

$$P(n,t|n',t') = \frac{1}{\sqrt{2\pi N\sigma^2}}e^{-\frac{[n-N\phi(t)]^2}{2\sigma^2}} \tag{32}$$

where, ϕ_t is given by equation (25). The distribution function $P(n,t|n',t')$ starting from an initial state (n',t') is always a time dependent function with the condition $\sigma\rangle 0$ (positive definite) which is also dependent on N.

A. Asymptotic state of $P(n,t|n',t')$

The asymptotic state of the probability distribution of the foraging ants can be obtained by taking at sufficiently large time i.e. by taking $\lim_{t\to\infty} P(n,t|n',t')$. In this limit, the expression for $\lim_{t\to\infty}\sigma^2$ can be obtained from the equation (31). Neglecting the value of $ln(0)$ which is undefined, we have the following expression for σ^2_∞ as given by,

$$\begin{aligned}\sigma^2_\infty &= \lim_{t\to\infty}\sigma^2 \\ &= \left(\frac{2b-c}{c-b}\right)\left[\frac{rb}{kc}+\frac{c^2}{2ka^2b^2}\{2r(kN-2r)+k(N+2)\}\right] \end{aligned} \tag{33}$$

The last factor in equation (33) is always positive because $kN\rangle 2r$ for large N. Hence, in order to have positive $\sigma^2_\infty\rangle 0$, the first factor has to be positive. In order to satisfy this condition, this inequality $\frac{2b-c}{c-b}\rangle 0$ must be satisfied. From this inequality, the condition for positivity of σ^2_∞ is given by, $2b\rangle c\rangle b$. Putting back the values of b and c from equation (26) to this inequality, we have,

$$\frac{N(1+k)-\phi_0}{2-\frac{1}{k}} > r > k\left[\frac{N}{2}-\phi_0\right] \tag{34}$$

Further, in this limit, we have from the equation (25) $\lim_{t\to\infty}\phi(t) = kN-2r$ which is a constant. Then, putting these expressions to the equation (35), the asymptotic distribution function is given by,

$$P(n,t|n',t') = \frac{1}{\sqrt{2\pi N\sigma^2_\infty}} e^{-\frac{[n-N(kN-2r)]^2}{2\sigma^2_\infty}} \tag{35}$$

We, then, calculate the mean number of ants of type $n = x_1$ at sufficiently large time using this probability distribution function (35). We calculated this $\langle n\rangle$ by taking $z = n - N(kN-2r)$ and integrating it which is given by,

$$\begin{aligned}\langle n\rangle &= \int_{-\infty}^{\infty} nP(n,t|n',t')dn \\ &= \frac{1}{\sqrt{2\pi N\sigma^2_\infty}}\int_{-\infty}^{\infty} ze^{-\frac{z^2}{2\sigma^2_\infty}}dz + \frac{N(kN-2r)}{\sqrt{2\pi N\sigma^2_\infty}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma^2_\infty}}dz \\ &= (kN-2r)\sqrt{N} \end{aligned} \tag{36}$$

The constant value of $\langle n \rangle$ indicates that at sufficiently long time, the particular $n = x_1$ type of ants will likely to forage with certain particular size of colony and similarly, the other type of the ants will also forage with another size of colony. Also, since the variance σ^2_∞ is a constant, the size of the colony of the ants at sufficiently long time is likely to be constant. The result (36) indicates that $\langle n \rangle$ depends on the choice of the constants r and k. In case $r < k$, $\langle n \rangle \sim kN\sqrt{N} \leq N$, which indicates that $k \leq \frac{1}{\sqrt{N}}$. The remaining ants $[N - \langle n \rangle]$ will be foraging in the second food source.

This simple analysis indicates that the formation of colonies by the foraging ants in search of food with random interaction among them is time dependent as indicated by time dependent variance $\sigma^2(t)$ as in equation (31) and hence time dependent $\langle n(t) \rangle$. Therefore, the size of the colonies of the respective ants in the ecosystem could be constantly changing with time. But at the sufficiently long time, the size of the colonies of the ants is likely to be constant as given $\langle n \rangle \to constant$ as $t \to \infty$ shown in equation (36). In this study, the rates r and k are taken to be time independent constants. However, in general these rates are time dependent and even the transition rates could be non-Markovian [19]. In such situation, the dynamics of foraging ants and subpopulation sizes of the x_1 and x_2 will be different and complicated with time dependent distributions.

IV. CONCLUSION

The foraging ants in a certain ecosystem is quite complicated process due to complicated information processing among them driven by individual decision. We studied a simple model of foraging ants [9] to understand the dynamics of single stochastic trajectory of the interacting ants using standard van Kampen's system size expansion [16] to solve the Master equation of the model we considered. This van Kampen's method assumes the transition probability to be proportional to the size of the jump and allows to express the Master equation to Kramers-Moyal series expansion in terms of extensive variables. This process allows the Master equation to become linear Fokker-Planck equation at thermodynamic limit which can be solved using proper transformation.

The probability distribution function of a certain type of ant, which is the solution of the simple foraging ants model using van Kampen's system size expansion, is quite complicated and is time dependent. The distribution function also indicates that the observables (mean, standard deviation and higher order cumulants), which can be calculated from this distribution function, are time dependent. This means the sizes of the colonies of the foraging ants are changing with time due to continuous interaction of the ants in the ecosystem. However, at sufficiently long time, this probability distribution function becomes time independent which is a constant. This indicates that the size of the colonies occupied by each type of ant in the ecosystem is likely to be constant because the mean and standard deviation are all become constants.

The analysis showed complicated results from the simple model. The interaction involved in the simple model is associated with complicated information processing which is driven by individual's decision during the process. This simple model can also be used to understand basics of the opinion formation during the public discourse etc. Even one can generalize it for multi-species interaction model. If we consider M-species (for example, ants subjected to identical M-food sources) system, then the 2-species model reactions (18) can be extended to M-species reactions. The probability distribution in the M-species Master equation will be M dimensional which will be difficult to solve. But one can solve it numerically by using Gillespie algorithm [12].

Acknowledgements

Author acknowledges financial assistance from the Matrics project, Department of Science and Technology, SERB with No. SERB/F/9606/2021-2022.

Competing interests

The author declare no competing interests.

References

[1] Murray J D. Mathematical Biology vols 1 and 2 (Berlin: Springer) (2002).

[2] Pasteels, J. M., Deneubourg, J. L., and Goss, S. Self-organization mechanisms in ant societies (I). Trail recruitment to newly discovered food sources. In Pasteels, J. M., and Deneubourg, J. L. (eds.), From hulividual to Collective Behavior in Social h sects, Birkhfiu- set, Basel, pp. t55-t76 (1987).

[3] C. Detrain and J.-L. Deneubourg, Self-organized structures in a superorganism: do ants "behave" like molecules?. Phys. Life Rev. 3, 162 (2006).

[4] A. Kirman. Ants, rationality, and recruitment. Q. J. Econ. 108, 137 (1993).

[5] .O. Richardson, E.J.H. Robinson, K. Christensen, H.J. Jensen, N.R. Franks, A.B. Sendova-Franks, Record dynamics in ants. PLoS ONE 5, e9621 (2010).

[6] Dussutour A, Beekman M, Nicolis SC, Meyer B. Noise improves collective decision-making by ants in dynamic environments. Proc. R. Soc. B 276, 353 (2009).

[7] Halley JD, Burd M. Non-equilibrium dynamics of social groups: insights from foraging Argentine ants. Ins Soc 51, 226 (2004).

[8] Jhawar J, Guttal V. Noise-induced effects in collective dynamics and inferring local interactions from data. Phil. Trans. R. Soc. B 375, 20190381 (2020).

[9] Biancalani T, Dyson L and McKane A J. Noise-induced bistable states and their mean switching time in foraging colonies. Phys. Rev. Lett. 112, 038101 (2014).

[10] N. G. van Kampen, Stochastic Processes in Physics and Chemistry (North-Holland, Amsterdam, 1981).

[11] C. W. Gardiner, Handbook of Stochastic Methods, 3rd ed. (Springer, New York, 2003).

[12] DT Gillespie, Exact stochastic simulation of coupled chemical reactions J. Phys. Chem. 81, 2340 (1977).

[13] Kurtz, T. G. The relationship between stochastic and deterministic models for chemical reactions. Journal of Chemical Physics, 57, 2976 (1972).

[14] Kramers H,.A. 1940. Physica, 7, 284 (Collected Scientific Papers (Amsterdam 1956), p. 754).

[15] Moyal J . E. Stochastic processes and statistical physics. J. Roy. Statist. Soc. (B) 11, 150 (1949).

[16] N. G. van Kampen. A power series expansion of the master equation. Can. J. Phys. 39, 551

(1961).

[17] Gillespie DT. Deterministic limit of stochastic chemical kinetics. J Phys Chem B 113, 1640 (2009).

[18] Risken, H. The Fokker–Planck Equation (Springer, 1989).

[19] N. G. van Kampen. Remarks on non-Markov processes, Braz. J. Phys. 28, 90 (1998).

Vector-borne transmission dynamics model based on Caputo fractional-order derivative

Kapil Kumar[1,2], **Mukesh Kumar Sharma**[2], **Sapna Ratan Shah** [*,3], and **Ravins Dohare** [†,4]

[1]Department of Mathematics, Atma Ram Sanatan Dharma College (University of Delhi), Dhaula Kuan, New Delhi, 110021, Delhi, India.
[2]Department of Mathematics, Chaudhary Charan Singh University, Meerut, 250004, Uttar Pradesh, India.
[3]School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, 110067, Delhi, India.
[4]Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, Jamia Nagar, New Delhi, 110025, Delhi, India.

[**Abstract:** This paper's major goal is to develop mathematical epidemiological models using fractional-order derivatives. Different epidemiological models in science and engineering have been effectively constructed and analysed because of fractional calculus's persuasive nature. Here, we presented a susceptible-exposed-infected-recovered (SEIR) structure based on Caputo fractional-order for hosts, and a susceptible-exposed-infected (SEI) structure based on Caputo fractional-order for mosquitoes, to study the dynamics of Zika and Dengue virus transmission.For this reason, we used the classical SEIR-SEI compartmental model to analyse the dynamics of the Zika fever outbreak that struck El Salvador in 2015–16. The updated model has an advantage over the traditional model since it considers the memory effect feature of nonlinear differential equations of fractional (non-integer) order. Additionally, we looked at the fractional-order model after the traditional approach using data from the ZIKV and DENV outbreaks that happened in Cape Verde in 2009 and Costa Rica in 2016–17, respectively. The effect of different fractions for the fractional-order based differential equation is also investigated. In the end, we came to the conclusion that the fractional-order model outperforms the equation-based classical model in terms of numerical findings.]

1 *Introduction*

In recent years, infectious diseases have become a significant threat worldwide. More than 335 infectious diseases were reported between 1940 and 2004 in the global human population [23]. Different microorganisms led to the birth of various infectious diseases, which had a profound effect on economies and world health. [22]. In the meantime, a number of infectious disease outbreaks in the past few decades, including Zika, Dengue, Chikungunya, West Nile, and Yellow fever, have been linked to vector-borne illnesses.

*Corresponding Author: sapnarshah@mail.jnu.ac.in
†Co-Corresponding Author: ravinsdohare@gmail.com

The matter garnered significant attention, and the threat posed by them is a cause for great concern [22]. Our primary focus was on Dengue Virus (DENV) and Zika Virus (ZIKV), two vector-borne illnesses. Since its discovery in Africa in 1952, ZIKV infection has been recognised as an emerging illness. A reported outbreak in Yap, Micronesia between April and July 2007 [11] was the primary source of ZIKV infection in humans.An outbreak that occurred in French Polynesia between October 2013 and April 2014 followed this. [7], as well as cases in other Pacific nations. A rapid mutation in the ZIKV in 2015 caused another outbreak in South America [31], and so on. The dengue virus, a flavivirus carried by mosquitoes, is the source of DENV infection. A dengue infection can cause a headache, nausea, joint discomfort, high fever, skin rash, and other symptoms. Usually, these signs appear three to fourteen days following the illness. Those infected with DENV typically need two to seven days to recover. Severe side effects from DENV infection can include low blood pressure, low blood platelet and plasma counts, and more.

A common female mosquito, *Ades aegypti* and *Ades albopictus*, is the vector of DENV and ZIKV illnesses. Due to the shared vector, infections and co-infections with these viruses have been documented in various parts of the world, including Asia [24]. Due to their similar clinical signs, it can occasionally be challenging to differentiate between the two infections. Nonetheless, there have been reports of ZIKV transmission through various routes, including blood, maternal-fetal, and sexual. ZIKV transmission through these channels is negligible in comparison to vector-borne transmission. As a result, we disregarded the possibility that ZIKV could spread through these channels in the current model and assumed that mosquitoes are the only source of infection. It is during the rainy and warmer seasons that these diseases are most likely to spread. [1]. Understanding the transmission behaviour of these diseases is crucial, which will help invade during disease transmission. Mathematical modelling is an aspect of understanding the transmission dynamics of diseases and predicting the progression of diseases. Much research has been done on the mathematical modelling of vector-borne diseases. Although, the most worrying aspect is that we are still unable to accurately predict the progression dynamics, which may be more far-reaching than we anticipated. Therefore, we need to improve our knowledge of the dynamics of vector-borne transmission by adding new methods to the current models. Creating strategies for disease prevention and control is crucial during an epidemic. For a very long time, mathematical modelling has been used to shed light on the global spread and management of numerous viral infectious illnesses. [4, 13, 18].

In order to study transmission dynamics, the majority of these investigations concentrated on deterministic integer-order compartmental modelling methodologies. Certain models with extra presumptions and restrictions have been put forth in an attempt to achieve the goal. These, however, are ill-suited to integrate host and vector memory with learning behaviour on the dynamics of vector-borne disease transmission. The memory effect trait can be used with fractional-order derivatives, as evidenced by recent studies on viral infections. which yields superior results to the traditional method of using integer-order derivatives. [10]. Even the traditional compartmental model based on integer-order derivatives was unable to accurately explain the disease's outbreak statistics data. Without any doubt, we could find several infectious disease progression models based on fractional-order derivatives in the literature [25, 32, 5]. Few studies are in literature on ZIKV and DENV modeling of disease progression dynamics based on fractional-order derivatives [28, 3, 12, 20]. Pooseh *et al.*[28] studied the SIR-SI fractional-order derivative-based modeling for dengue infection progression with Riemann-Liouville-type derivatives. To get an approximation of the original fractional system's solution, they

employed traditional techniques. Through numerical analysis, scientists determined that the ideal fractional order is 0.987, based on data from an outbreak that occurred in 2009 in Cape Verde. This is due to the fact that, in contrast to the sixty-two in the classical model, the percentage inaccuracy is thirteen. The researchers utilised two alternative fractional orders and fractional Caputo derivatives in 2013 based on the assumption that human behaviour differs from that of mosquitoes [10]. Li *et al.*[20] presented a novel and all-encompassing fractional-order dengue fever system in 2019 that makes use of Caputo derivatives in various orders. The findings demonstrated that, in comparison to other models, The numerical solutions with the predicted parameter values for the multi-term fractional-order dengue fever model suit the data better. Alkahtani *et al.* [3] developed a mathematical model for the progression of the Zika virus. Both the equilibrium point and the reproductive number were shown here. Using the Adams type predictor-corrector rule for the Atangana-Baleanu fractional integral, the model was analytically solved. Under certain conditions, the existence and unique exact solution were shown. A fractional Zika virus model with mutation was constructed and presented in a paper [2]. The virus creates new birth issues in infected pregnant women and spreads further in society. The model is simulated on actual Colombia data. They described the proposed model's equilibrium points, invariant property, and positivity of the model solution. The computation of the reproduction number is done using a next-generation technique. A new mathematical model explaining how the Zika virus spreads from people to mosquitoes was provided by [29]. The stability of the equilibrium point was examined, The feasibility and equilibrium points of the system were calculated, and they employed the fractional-order Caputo derivative. Using the fixed point theory, they also demonstrated that the model has a single, unique solution. Few more studies have been done on ZIKV with fractional order derivative-based modelling, but none of the studies compared classical and fractional-based modelling for ZIKV with the support of simulation on real-time outbreak data.

The current study expands on the traditional SEIR-SEI mathematical model, which has previously been used to analyse the dynamics of the El-Salvador outbreak of ZIKV sickness in 2015–16. [16] into Caputo type fractional-order derivative-based SEIR-SEI model. Here, we considered different fraction values for both virus and host populations (mosquito and human) in this model. Both models were solved simultaneously; the classical SEIR-SEI model was solved using Runge-Kutta fourth-order numerical methods, whereas the fractional-order based model was solved using Adams-Bashforth-Moultan numerical algorithm [19]. During the numerical solution, the disease parameter values have been retrieved from the actual outbreak of ZIKV and DENV data of Costa Rica during 2016-17 and Cape Verde during 2009, respectively. The fractional-order-based model's solution was best fitted with actual outbreak data provided that value of R^2 are 0.8355 and 0.8834 quantified for ZIKV and DENV data respectively. In contrast, the solution of the classical SEIR-SEI model was poorly fitted with the actual outbreak data and R^2 quantified 0.6681 and 0.7014 for ZIKV and DENV respectively. This fractional-order-based modelling is due to the memory effect property. Every state of fractional-order based modelling depends on all the previous states, although classical integer-order based current state is only based on the first initial state. Hence, this Caputo type fractional-order derivative mathematical modelling captures epidemic with high accuracy. Consequently, the enumeration of accurately basic reproduction numbers was impactful. Estimating disease parameters and basic reproduction numbers have a leading role in the control and progression of the disease. Henceforth,According to this study, vector-borne systems may benefit more from fractional-order derivative modelling than from integer-order modelling. This modelling approach may also be applicable to other viral illnesses such as COVID-19, influenza, etc.

2 *Methods*

2.1 *The classical mathematical model*

In order to track the development of both ZIKV and DENV infection, this work makes use of the SEIR compartmental model for hosts and the SEI model for vectors. According to this idea, the infectious status of each individual determines the division of the human population into distinct classes. Every class depicts a person's state of health at infection time t. We hypothesised that the populations of the vector (mosquito) and host (human) would mix uniformly. Every human and mosquito in their respective populations has the same chance of becoming infected and spreading the sickness. While the mosquito population is separated into three groups—susceptible (S_v), exposed (E_v), and infected (I_v)—the human population is classified into four mutually exclusive groups: susceptible (S_h), exposed (E_h), infected (I_h), and recovered (R_h). $N_h(t)$ represents the entire population of humans, while $N_v(t)$ represents the population of mosquitoes. In this case, $N_h(t) = S_h(t) + E_h(t) + I_h(t) + R_h(t)$ and $N_v(t) = S_v(t) + E_v(t) + I_v(t)$. The same model, which is described by a system of non-linear Ordinary Differential Equations provided in the system (1), was taken into consideration for ZIKV and DENV. The model is as follows:

$$\begin{aligned}
\frac{dS_h}{dt} &= -\alpha_h S_h I_v, \\
\frac{dE_h}{dt} &= \alpha_h S_h I_v - \beta_h E_h, \\
\frac{dI_h}{dt} &= \beta_h E_h - \gamma_h I_h, \\
\frac{dR_h}{dt} &= \gamma_h I_h, \\
\frac{dS'_v}{dt} &= \mu_v - \mu_v S'_v - \alpha_v S'_v \frac{I_h}{N_h}, \\
\frac{dE'_v}{dt} &= -\mu_v E'_v - \beta_v E'_v + \alpha_v S'_v \frac{I_h}{N_h}, \\
\frac{dI'_v}{dt} &= \beta_v E'_v - \mu_v I'_v,
\end{aligned} \tag{1}$$

where, $S'_v(t)$, E'_v and I'_v signifies proportion of vectors with the property $0 \leq S'_v, E'_v, I'_v \leq 1$ given in equation (2).

$$S'_v(t) = \frac{S_v}{N_v}, \quad E'_v(t) = \frac{E_v}{N_v}, \quad I'_v(t) = \frac{I_v}{N_v}. \tag{2}$$

In this system, α_h denotes the rate at which susceptible humans are exposed and is equal to $a \times b \times m$, where a is the number of mosquitoes bites per day per human, b is the transmission probability from infectious mosquito to suspected human per bite and m is the average ratio of mosquito to human, β_h is the rate at which exposed humans are become infected, γ_h is the rate at which infected humans get recovered. Similarly, α_v is the rate at which susceptible mosquitoes are exposed and is equal to $a \times c$ (independent of vector to human ratio), where c is the transmission probabilities from an infected human to infected mosquitoes, β_v is the rate at which exposed mosquitoes become infected. Moreover, the parameter μ_v is the mortality rate of vectors (mosquito life span). Finally, the rate from the infected human population to the recovered human

population is denoted by γ_h. For a better understanding of parameters, the above-mentioned model (1) can be defined accordingly as in a system of equations (3).

$$\begin{aligned}
\frac{dS_h}{dt} &= -abmS_hI_v, \\
\frac{dE_h}{dt} &= abmS_hI_v - \beta_hE_h, \\
\frac{dI_h}{dt} &= \beta_hE_h - \gamma_hI_h, \\
\frac{dR_h}{dt} &= \gamma_hI_h, \\
\frac{dS'_v}{dt} &= \mu_v - \mu_vS'_v - acS'_v\frac{I_h}{N_h}, \\
\frac{dE'_v}{dt} &= -\mu_vE'_v - \beta_vE'_v + acS'_v\frac{I_h}{N_h}, \\
\frac{dI'_v}{dt} &= \beta_vE'_v - \mu_vI'_v,
\end{aligned} \tag{3}$$

where, N_h=$S_h + E_h + I_h + R_h$ and N'_v=$S'_v + E'_v + I'_h$.
While creating this model, we make the following assumptions: (a) The transmission does not take into account human-to-human transmission such as sexual transmission, mother-to-child transmission, or blood transmission; (b) The total human population (N_h) is kept constant because demographic changes, migration, new births, or deaths are not taken into account in the human population, and (c) The total vector population (N_v) is likewise kept constant throughout the work by taking into account the same rates of birth and death.

2.2 *Preliminaries of fractional derivatives*

Although fractional derivative equations (FDEs) are not new, there are novel applications in several fields of study. In a 1695 letter, Leibniz foresaw the idea of FDEs [21]. Consequently, Abel was the first to investigate tautochrone difficulties indirectly using these FDEs in 1823. Following that, a number of foundational papers based on FDEs on different aspects were published [6]. While there are numerous intriguing definitions of fractional derivatives, the Riemann-Liouville concept is the most well-known [14]. The definition of the order α Riemann-Liouville derivative is:

$$D^{\alpha}_{0+}f(t) = \frac{1}{\Gamma(n-\alpha)}(\frac{d}{dt})^n\int_0^t\frac{f(s)}{(t-s)^{\alpha-n+1}}ds, \quad n=[\alpha]+1, \tag{4}$$

where, $0 < \alpha < 1$ for $\alpha \in R$, n is an integer, Γ represents the gamma function and $[\alpha]$ represents greatest integer value of α. This approach leads to the following two issues. First, Riemann-Liouville does not hold differentiation of the constant [9], i.e.,

$$D^{\alpha}c = \frac{c}{\Gamma(1-\alpha)}t^{-\alpha} \neq 0, c = constant. \tag{5}$$

Therefore, if the constant value of the differential operators is substituted by a Riemann-Liouville differential operator of order $\alpha \in (0,1)$, it is unable to solve the differential operators [6]. As a result, our assumption that the total number of human (N_h) and the total number of human (N_v) should remain constant will no longer apply to the changed model. The failure to combine with the initial conditions of the form defined above in the classical model [[9], pp. 54–55] is the second problem with fractional derivatives of

the Riemann-Liouville type.

Consequently, rather than using Riemann-Liouville, we took into account the Caputo type fractional derivative in this study [[9], Chap. 3]. It is described as:

$$D_t^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t \frac{f^n(s)}{(t-s)^{\alpha-n+1}} ds, \quad n = [\alpha] + 1. \tag{6}$$

2.3 *Fractional order derivative based model*

FDEs have received a lot of interest lately from a variety of fields, including biology, physics, engineering, and the biological sciences. It has been noted that compared to classical models based on integer order derivatives, fractional order derivatives can offer a greater agreement between real and simulated data [10, 9]. Consequently, the provided model makes sense when expressed in terms of fractional order derivatives, which has a benefit because of its memory effect property. According to the memory effect property, a model's future state may be influenced by both its current and previous states.
Consequently, the updated model (7) is obtained by substituting the Caputo derivative of order α for the integer derivative in model (3):

$$\begin{aligned}
\frac{d^{\alpha} S_h}{d^{\alpha} t} &= -abmS_h I_v, \\
\frac{d^{\alpha} E_h}{d^{\alpha} t} &= abmS_h I_v - \beta_h E_h, \\
\frac{d^{\alpha} I_h}{d^{\alpha} t} &= \beta_h E_h - \gamma_h I_h, \\
\frac{d^{\alpha} R_h}{d^{\alpha} t} &= \gamma_h I_h, \\
\frac{d^{\alpha} S'_v}{d^{\alpha} t} &= \mu_v - \mu_v S'_v - acS'_v \frac{I_h}{N_h}, \\
\frac{d^{\alpha} E'_v}{d^{\alpha} t} &= -\mu_v E'_v - \beta_v E'_v + acS'_v \frac{I_h}{N_h}, \\
\frac{d^{\alpha} I'_v}{d^{\alpha} t} &= \beta_v E'_v - \mu_v I'_v,
\end{aligned} \tag{7}$$

where $\alpha \in (0, 1]$ is the order of fractional derivative and the fractional operator d^{α} is identical to the classical first derivative for $\alpha = 1$.

Additionally, we can see from a basic dimension analysis that the left and right sides of the equations for model (7) have mismatched dimensions. This model's left side has dimension $(time)^{-\alpha}$, according to a basic dimensional analysis. The terms $a, \beta_h, \gamma_h, \mu_v, \beta_v$ on the right-hand side have the dimension $(time)^{-1}$; a closer examination reveals that the remaining terms on the right-hand side are dimensionless. As a result, the dimension of this model's right side is $(time)^{-1}$. We guarantee that the system's equations on the right side have the same dimensions as those on the left. Diethelm [10] developed a fractional order deterministic model for dengue disease in order to solve this issue of fractionalization. Sardar (2015) [30] *et al.* used this method to analyse a similar model and achieved the better results. Consequently, we attempted to address the discrepancy in dimensions by applying the method that Diethelm [10] outlined. Furthermore, this model is further refined by the fact that human population behaviour differs from that of mosquitoes. Therefore, two distinct fractional differential operators $\alpha \in (0, 1]$ and $\beta \in (0, 1]$ for humans and mosquitoes, respectively, can be added to the model (7) to make it more realistic. The refined and updated model, denoted as (8), is

as follows:

$$\begin{aligned}
\frac{d^\alpha S_h}{d^\alpha t} &= -a^\alpha bmS_hI_v, \\
\frac{d^\alpha E_h}{d^\alpha t} &= a^\alpha(b)(m)S_hI_v - \beta_h^\alpha E_h, \\
\frac{d^\alpha I_h}{d^\alpha t} &= \beta_h^\alpha E_h - \gamma_h^\alpha I_h, \\
\frac{d^\alpha R_h}{d^\alpha t} &= \gamma_h^\alpha I_h, \\
\frac{d^\beta S'_v}{d^\beta t} &= \mu_v^\beta - \mu_v^\beta S'_v - (a)^\beta(c)S'_v\frac{I_h}{N_h}, \\
\frac{d^\beta E'_v}{d^\beta t} &= -\mu_v^\beta E'_v - \beta_v^\beta E'_v + (a)^\beta(c)S'_v\frac{I_h}{N_h}, \\
\frac{d^\beta I'_v}{d^\beta t} &= \beta_v^\beta E'_v - \mu_v^\beta I'_v.
\end{aligned} \tag{8}$$

It is clear that the model (8) reduces to the classical model (3) as $\alpha \to 1$ and $\beta \to 1$. It is found that the total population N_h and N_v in the redesigned system are constant by a straightforward computation in model (8). Consequently, the fractional order initial value problem's mathematical formulation is covered in this section. The classical model (3) is a deterministic integer order system where no information from the previous state is carried over to the current state. In biological science, visco-elastic studies, etc., it is well-established that the condition of many systems depends on the characteristics of earlier times [27]. Understanding the memory and learning behaviours of hosts and vectors is essential to comprehending the dynamics of mosquito-borne illness transmission. In order to obtain better and more precise findings, we thus attempted to generalise the classical model, which contains information about its many prior states.

Furthermore, understanding the dynamics of infections requires an understanding of the Reproduction number. The predicted numbers of secondary infections caused by single infections in a population where all individuals are susceptible is known as the reproduction number (R_0) [8]. It is employed to calculate a disease's propensity for transmission. Generally speaking, the fundamental reproduction number (R_0) needs to be larger than one in order for illnesses to spread across vulnerable populations. Using the techniques from [17], the basic reproduction number for the fractional order model (8) has been determined as follows:

$$R_0 = \sqrt{\frac{(a^\alpha bm)(a^\alpha c)\beta_v^\beta}{\gamma_h^\alpha \mu_v^\beta(\mu_v^\beta + \beta_v^\beta)}}. \tag{9}$$

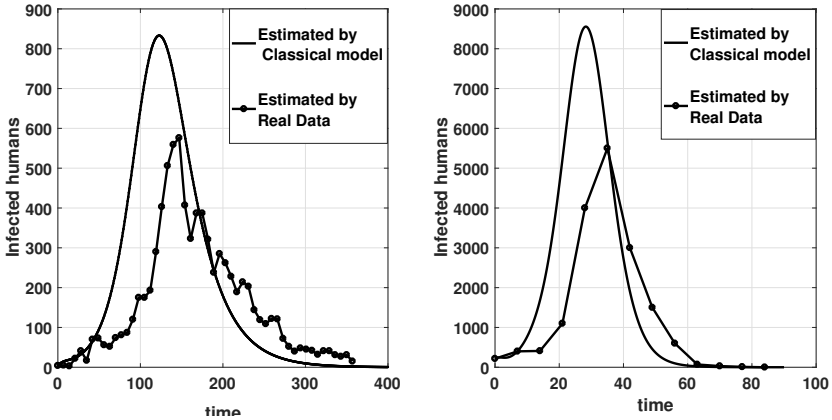
R^2 is a coefficient of determination which provide us how best fit predicted data with the actual outbreak data. The calculation of R^2 can be obtained follows as:

$$SSE = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{10}$$

$$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2, \tag{11}$$

$$R^2 = \frac{SSE}{SSR + SSE}, \tag{12}$$

where $\hat{y}_i$ and y_i are i^{th} predicted and corresponding actual outbreak value respectively. $\bar{y}$ is mean of actual outbreak values.



(a) Zika outbreak, Costa Rica 2016-17 (b) Dengue outbreak, Cape Verde 2009

Figure 1: Outbreak dynamics and classical model estimated dynamics by eq. 3 of infected human $I_h(t)$ in the (a) Costa Rica 2016-17 Zika outbreak, (b) Cape Verde 2009 Dengue outbreak.

2.4 *Simulations*

The Pan American Health Association (PAHO) from 17^{th} week of 2016 to 15^{th} week of 2017 and the government of Cape Verde, respectively, provided the data for two outbreaks that happened in Costa Rica in 2016 for ZIKV and another one in Cape Verde Island in 2009 for DENV disease [26, 10]. Since the mathematical model has parameters, which we know exist, the parameters can be obtained in the literature by using a particular computation or by using the original data. Here, Table 1 displays the parameter ranges that were found in the relevant literature. Table 2 provides the beginning conditions for the parameters for both sets of data. Through the use of Runge-Kutta fourth order, we were able to simulate the classical model (3) on the MATLAB platform. The results we obtained are displayed in Figures 1a and 1b, respectively, as a plot between the simulated model and the actual number of reported cases. The two curves in both figures were compared, and the results indicated that there was a poor fit between the outbreak's real data and the data produced by the model using the specified parameter values. The model's estimate of infected persons was a little bit high. Therefore, we employ the generalisation of the model to obtain better and more accurate findings.

3 *Results*

By building the model with memory in both the host and vector population, the fractional order model (8) is used to mimic the ZIKV and DENV transmission as a generalisation from a classical model (3). To our knowledge, there isn't an analytical solution available for solving and analysing the FDEs. We used an effective predictor-corrector numerical technique, briefly reported in [6, 30], to numerically solve the fractional order problem presented in model (8). In the Caputo fractional derivative, zero is considered the beginning point. Since adopting lesser step sizes, such as 10^{-4} days, does not materially alter the findings, 10^{-2} days has been chosen as the step size for this system of the equation [10]. We looked for the fractional order that best fits the data on the reported number of cases using the parameter values and starting values listed in Tables 1 and 2.

Table 1: Baseline values and their ranges of parameters used in the models.

Parameters	**Baseline values** (Costa Rica)	**Baseline values** (Cape Verde)	**Ranges**
a	0.3	0.7	[0.3,1]
b	0.4	0.4	[0.1,0.75]
m	5	5	[1,10]
c	0.6	0.5	[0.3,0.75]
$1/\gamma_h$	8	3	[3,14]
β_h	0.03	0.36	[0.1,0.75]
β_v	0.1	0.36	[0.1,0.75]
$1/\mu_v$	17	10	[4,35]

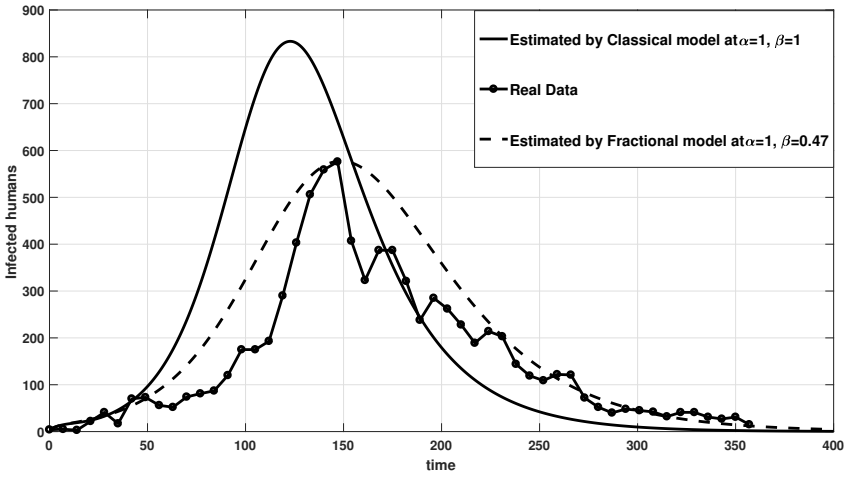
Table 2: Opted initial values for the for the simulation of model.

Variable	**Costa Rica** (Zika)	**Cape Verde** (Dengue)
S_h	10000	55784
E_h	100	300
I_h	4	216
R_h	0	0
S_v	100000	168000
E_v	20	20
I_v	10	10

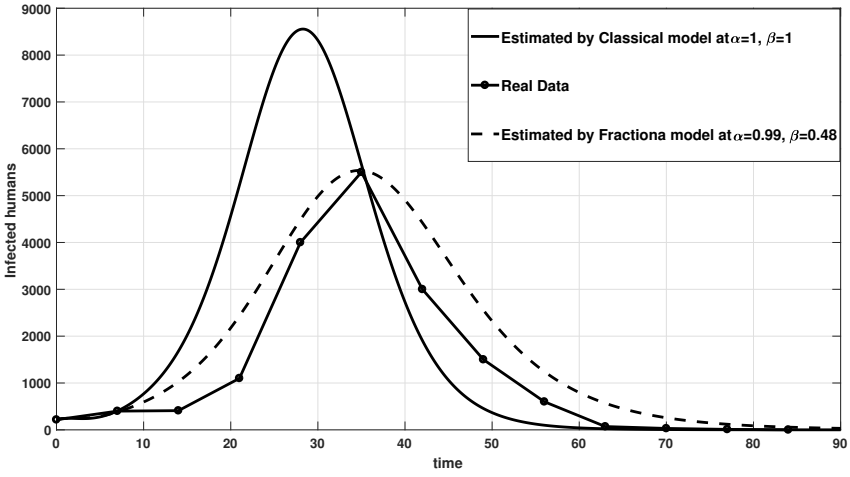
respectively. A search on the interval (0, 1] yielded the derivative order value, or $\alpha, \beta \in$ i/100: i=1,2,3,.....,100. We began with $\alpha = 1$ and decreased it step-by-step until we arrived at an optimal number. For the data from Costa Rica ($\alpha = 1$ and $\beta = 0.47$) and Cape Verde ($\alpha = 0.99$ and $\beta = 0.48$), a satisfactory approximation was found. Figures 2a and 2b display the plot of numerical solutions of the system of FDEs with fixed parameter values. Compared to the standard solution ($\alpha, \beta = 1$), these figures showed that variations in the values of α and β can result in a more accurate approximation of actual outbreak data. Compared to greater values of α, it is evident that smaller values of α result in somewhat better approximations of the real data during the early stages of the epidemic.

Moreover, we also calculated the basic reproduction number for the fractional order model using the equation (9). For Cost Rica outbreak, R_0 comes out to be 1.86558 for $\alpha = 1$, $\beta = 0.47$ and $R_0 = 1.6286$ for the Cape Verde with the values $\alpha = 0.99$, $\beta = 0.48$. Both the values of R_0 come out to be > 1, indicating that the Zika and Dengue virus is spreading in both regions. Also, through the previous studies of disease outbreaks, R_0 is found to be greater i.e. $0.5 - 6.3$ in El-Salvador, Brazil and Columbia [16] and $2.6 - 4.8$ in French Polynesia [15].

The behaviour of Basic Reproduction Number R_0 with the variation in different parameters defined for the epidemic model is shown in Figure 3a-3f for both Costa Rica (Left panel) and Cape Verde (Right Panel) outbreaks of Zika and Dengue virus. Figure 3a and Figure 3b, depicted that R_0 varies from 0-20 with respect to change in mosquitoes biting rate $a \in [0, 5]$ for different value of recovery rate i.e. $\gamma_h \in [0.0714, 0.3333]$ for



(a) Zika outbreak, Costa Rica 2016-17



(b) Dengue outbreak, Cape Verde 2009

Figure 2: Infectious dynamics of real outbreak data, estimated by classical model (i.e. $\alpha = 1$) and (a) estimated by the Caputo fractional model with $\alpha = 1$ and $\beta = 0.47$ for Costa Rica 2016-17 Zika outbreak, (b) estimated by the fractional model with $\alpha = 0.99$ and $\beta = 0.48$ for Cape Verde 2009 Dengue outbreak.

Costa Rica and Cape Verde epidemic regions, respectively. It can be noticed through the figure that R_0 is directly proportional to mosquitoes biting rate with any value of γ_h. Further, we can notice that R_0 also depends on γ_h (recovery rate) as it increases with a decrease in the value of γ_h at any fixed value of a (mosquito's biting rate). Figure 3c and Figure 3d, illustrated inversely proportionate behavior of Reproduction number (R_0) to the death rate of mosquitoes (μ_v) with the change in values of mosquitoes biting rate, $a \in [0.3, 1]$. The range of R_0 comes out to be $[0.5, 6]$ for both the data with respect to the μ_v. Here, the relationship curve also indicates the upward shift with the increase in the mosquito's biting rate value. Moreover, Figure 3e and 3f, also showed that R_0 is in direct relationship with the mosquitoes biting rate with variation in the value of μ_v, but the curve showed the reciprocal behaviour with a death rate of mosquitoes μ_v as the value of R_0 decreases with the increase in value of μ_v.

We also simulated S_h, R_h and I_h for different values of fractional order for concern data sets. Three different values of α and β are considered in each plot. When $\alpha = 1$, the system is in the classical order. The plots Fig. 4a-4f depicted that variation in $S_h(t)$, $R_h(t)$ and $I_h(t)$ versus time t for different values of α & $(=)$ $\beta = 1, 0.9, 0.8$ by fixing the other parameters value. With the same set of parameters, it is noticed that the susceptible population decreases dramatically in a comparatively short amount of time, which appears to be somewhat unrealistic (Fig. 4a-4b). These graphs demonstrate that utilising an approximate solution with a basic fractional model will yield surprisingly superior outcomes. A classical model that is converted to a fractional one, however, becomes sensitive to the order of differentiation since even a little variation in the fractional order value affects the outcome significantly. Figs. 4c-4d make it evident that the fractional derivatives α and β continuously influence the approximate solutions. As a result, the results demonstrate that by using fractional derivatives, we were able to enhance the dynamics of the SEIR model.

4 *Discussion*

Dengue and Zika are still illnesses that could pose a risk to people all across the world. By using fractional order, this research is thought to be a generalisation of a traditional epidemic model. We developed and examined a ZIKV and DENV model based on fractional order derivatives infection progression during the Costa Rica 2016-17 and Cape Verde 2009 outbreaks, respectively. Additionally, we assessed the model using various fractional operator values and noted its impact at various levels utilising the same parameters as initially stated in the classical model. It is determined that a non-linear FDEs model yields a more accurate outcome by simulating infection using the non-linear fractional model. Here, accuracy could be achieved by tuning a single variable fractional operator, while integer-based modelling needs to tune several parameters. We quantitatively showed that the fractional order model yields more exciting and accurate results, using the Adams-type predictor-corrector method. This model assumes that the population is homogeneously distributed; this assumption might be expanded in the form of an in-homogeneously distributed population during the development of fractional order modelling. This in-homogeneity attempt has been made using network-based modelling. This approach might be the future scope for this modelling.

5 *Conclusion*

We developed and examined a model for the evolution of dengue and ZIKV infections during the outbreak in various cities, based on fractional-order derivatives. Additionally,

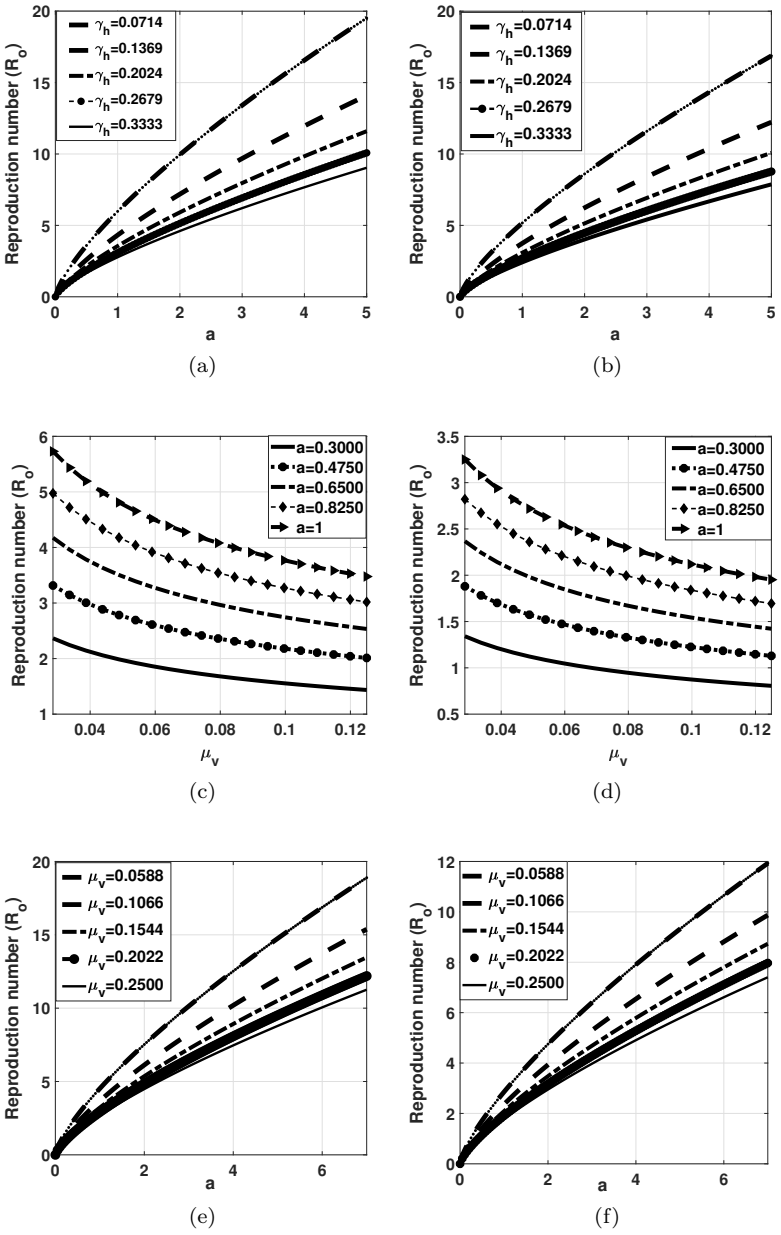


Figure 3: Basic reproduction number (R_0) varying with crucial parameters involved in Caputo model, plot (a) R_0 versus mosquitoes biting rate a, with varying γ_h for Costa Rica, (b) R_0 versus mosquitoes biting rate a, with varying γ_h for Cape Verde, (c) R_0 versus μ_v with varying mosquitoes biting rate a for Costa Rica, (d) R_0 versus μ_v, with varying mosquitoes biting rate a for Cape Verde, (e) R_0 versus mosquitoes biting rate, a with variation μ_v for Costa Rica, (f) R_0 versus mosquitoes biting rate, a with variation μ_v for Cape Verde.

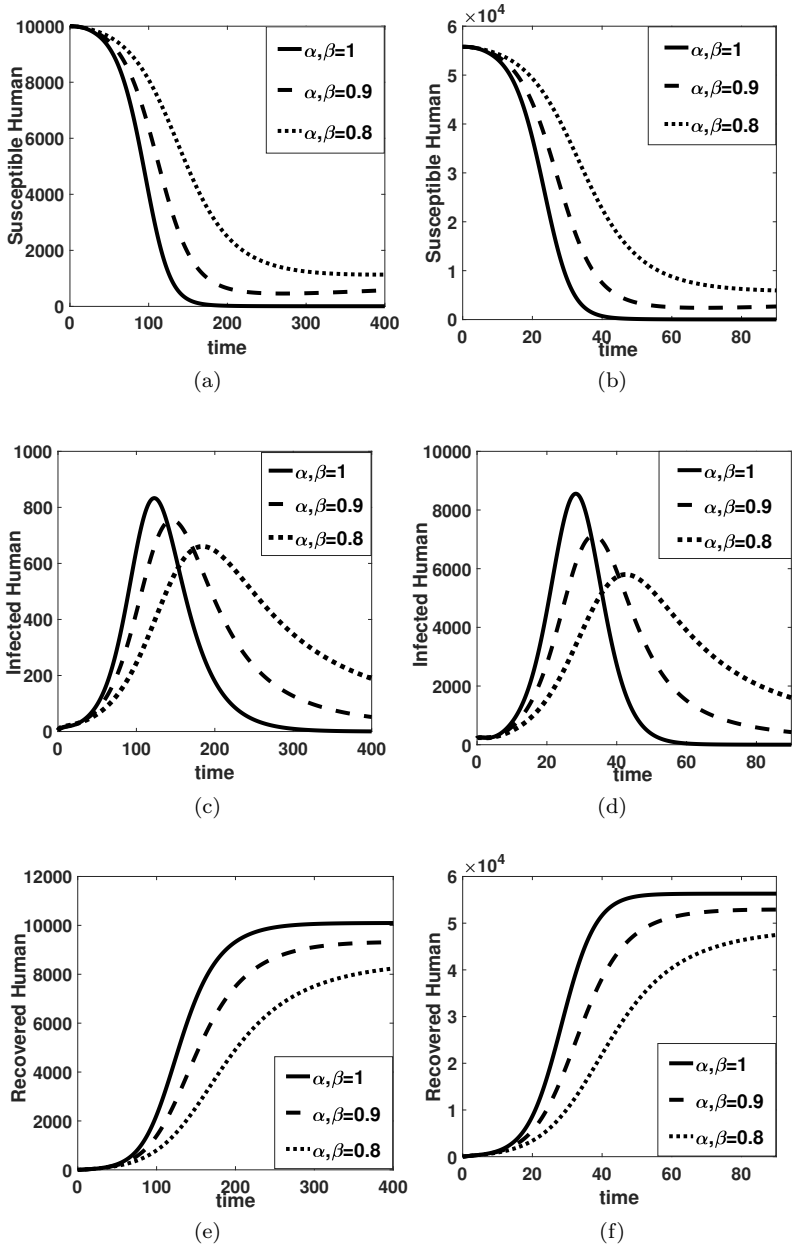


Figure 4: Numerical solution of Caputo model (8) at different fractions values with other parameters are as Table 1. (a) $S_h(t)$ of Costa Rica 2016-17 Zika outbreak at α and $\beta = 1, 0.9, 0.8$, (b) $S_h(t)$ of Cape Verde 2009 Dengue outbreak at α and $\beta = 1, 0.9, 0.8$, (c) $I_h(t)$ of Costa Rica 2016-17 Zika outbreak at α and $\beta = 1, 0.9, 0.8$, (d) $I_h(t)$ of Cape Verde 2009 Dengue outbreak at α and $\beta = 1, 0.9, 0.8$, (e) $R_h(t)$ of Costa Rica Zika 2016-17 outbreak at α and $\beta = 1, 0.9, 0.8$, (f) $R_h(t)$ of Cape Verde 2009 Dengue outbreak at α and $\beta = 1, 0.9, 0.8$.

we assessed the model using various fractional operator values and noted its impact at various levels utilising the same parameters as initially stated in the classical model. It is determined that a non-linear FDEs model yields more accurate findings than an ODEs model by simulating infection using both linear and non-linear fractional models.Ultimately, we came to the conclusion that the fractional-order model outperforms the equation-based classical model in terms of numerical findings.

Abbreviations

SEIR: Susceptible-exposed-infected-recovered
SEI: Susceptible-exposed-infected
ZIKV: Zika Virus
DENV: Dengue Virus
FDE: Fractional differential equation

Acknowledgments

We would like to offer my special thanks to Dr Narender Kumar, who contributed sufficiently to this work, although he is no longer with us. Personally, the authors are praying about him: May God gives him the best place in heaven or wherever he is. We would also like to thank SERB (with grant no. EEQ/2023/000980) for providing a fund for basic infrastructure in the lab.

References

[1] Website: World Health Organization (WHO). https://www.who.int/, 2017.

[2] Aatif Ali, Saeed Islam, M Riaz Khan, Saim Rasheed, FM Allehiany, Jamel Baili, Muhammad Altaf Khan, and Hijaz Ahmad. Dynamics of a fractional order zika virus model with mutant. *Alexandria Engineering Journal*, 2021.

[3] Badr Saad T Alkahtani, Abdon Atangana, and Ilknur Koca. Novel analysis of the fractional zika model using the adams type predictor-corrector rule for non-singular and non-local fractional operators. *J. Nonlinear Sci. Appl*, 10(6):3191–3200, 2017.

[4] Roy M Anderson. Directly transmitted viral and bacterial infections of man. In *The Population Dynamics of Infectious Diseases: Theory and Applications*, pages 1–37. Springer, 1982.

[5] Abdon Atangana. Modelling the spread of covid-19 with new fractal-fractional operators: can the lockdown save mankind before vaccination? *Chaos, Solitons & Fractals*, 136:109860, 2020.

[6] Dumitru Baleanu, Kai Diethelm, Enrico Scalas, and Juan J Trujillo. *Fractional calculus: models and numerical methods*, volume 3. World Scientific, 2012.

[7] Van-Mai Cao-Lormeau, Claudine Roche, Anita Teissier, Emilie Robin, Anne-Laure Berry, Henri-Pierre Mallet, Amadou Alpha Sall, and Didier Musso. Zika virus, french polynesia, south pacific, 2013. *Emerging infectious diseases*, 20(6):1085, 2014.

[8] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[9] Kai Diethelm. *The analysis of fractional differential equations: An application-oriented exposition using differential operators of Caputo type.* Springer Science & Business Media, 2010.

[10] Kai Diethelm. A fractional calculus based model for the simulation of an outbreak of dengue fever. *Nonlinear Dynamics*, 71(4):613–619, 2013.

[11] Mark R Duffy, Tai-Ho Chen, W Thane Hancock, Ann M Powers, Jacob L Kool, Robert S Lanciotti, Moses Pretrick, Maria Marfel, Stacey Holzbauer, Christine Dubray, et al. Zika virus outbreak on yap island, federated states of micronesia. *New England Journal of Medicine*, 360(24):2536–2543, 2009.

[12] H Elsaka and E Ahmed. A fractional order network model for zika. *BioRxiv*, page 039917, 2016.

[13] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

[14] A Anatolii Aleksandrovich Kilbas, Hari Mohan Srivastava, and Juan J Trujillo. *Theory and applications of fractional differential equations*, volume 204. Elsevier Science Limited, 2006.

[15] Adam J Kucharski, Sebastian Funk, Rosalind M Eggo, Henri-Pierre Mallet, W John Edmunds, and Eric J Nilles. Transmission dynamics of zika virus in island populations: a modelling analysis of the 2013–14 french polynesia outbreak. *PLoS neglected tropical diseases*, 10(5):e0004726, 2016.

[16] Narender Kumar, Mohd Abdullah, Md Imam Faizan, Anwar Ahmed, Hytham A Alsenaidy, Ravins Dohare, and Shama Parveen. Progression dynamics of zika fever outbreak in el salvador during 2015–2016: a mathematical modeling approach. *Future Virology*, 12(5):271–281, 2017.

[17] Narender Kumar, Shama Parveen, and Ravins Dohare. Comparative transmission dynamics and optimal controls for chikungunya, dengue and zika virus infections. *Letters in Biomathematics*, 6(2):1–14, 2019.

[18] Sachin Kumar and Shikha Jain. Assessing the effects of treatment in hiv-tb co-infection model. *The European Physical Journal Plus*, 133(8):1–20, 2018.

[19] Changpin Li and Chunxing Tao. On the fractional adams method. *Computers & Mathematics with Applications*, 58(8):1573–1588, 2009.

[20] Tianzeng Li, Yu Wang, Fawang Liu, and I Turner. Novel parameter estimation techniques for a multi-term fractional dynamical epidemic model of dengue fever. *Numerical Algorithms*, 82(4):1467–1495, 2019.

[21] JA Tenreiro Machado, Alexandra MSF Galhano, and Juan J Trujillo. On development of fractional calculus during the last fifty years. *Scientometrics*, 98(1):577–582, 2014.

[22] David M Morens, Gregory K Folkers, and Anthony S Fauci. The challenge of emerging and re-emerging infectious diseases. *Nature*, 430(6996):242, 2004.

[23] Stephen S Morse. Factors in the emergence of infectious diseases. In *Plagues and politics*, pages 8–26. Springer, 2001.

[24] Kamaldeen O Okuneye, Jorge X Velasco-Hernandez, and Abba B Gumel. The "unholy" chikungunya–dengue–zika trinity: A theoretical analysis. *Journal of Biological Systems*, 25(04):545–585, 2017.

[25] Kolade M Owolabi. Mathematical modelling and analysis of two-component system with caputo fractional derivative order. *Chaos, Solitons & Fractals*, 103:544–554, 2017.

[26] PAHO. Paho/who data- plisa, http://www.paho.org/data/index.php/en/.

[27] Ivo Petráš and Richard L Magin. Simulation of drug uptake in a two compartmental fractional model for a biological system. *Communications in Nonlinear Science and Numerical Simulation*, 16(12):4588–4595, 2011.

[28] Shakoor Pooseh, Helena Sofia Rodrigues, and Delfim FM Torres. Fractional derivatives in dengue epidemics. In *AIP Conference Proceedings*, volume 1389, pages 739–742. AIP, 2011.

[29] Shahram Rezapour, Hakimeh Mohammadi, and Amin Jajarmi. A new mathematical model for zika virus transmission. *Advances in Difference Equations*, 2020(1):1–15, 2020.

[30] Tridip Sardar, Sourav Rana, Sabyasachi Bhattacharya, Kamel Al-Khaled, and Joydev Chattopadhyay. A generic model for a single strain mosquito-transmitted disease with memory on the host and the vector. *Mathematical biosciences*, 263:18–36, 2015.

[31] Lavinia Schuler-Faccini. Possible association between zika virus infection and microcephaly—brazil, 2015. *MMWR. Morbidity and mortality weekly report*, 65, 2016.

[32] Saif Ullah, Muhammad Altaf Khan, and Muhammad Farooq. A fractional model for the dynamics of tb virus. *Chaos, Solitons & Fractals*, 116:63–71, 2018.

INFORMATION TO AUTHORS

Manuscripts should represent results of original works on theoretical physics or experimental physics with theoretical background or on applied mathematics and topics of interdisciplinary nature. Letters to the Editor and Review articles in emerging areas are also published. Submission of the manuscript will be deemed to imply that it has not been published previously and is not under consideration for publication elsewhere (either partly or wholly) and further that, if accepted, it will not be published elsewhere. It is the right of the Editorial Board to accept or to reject the paper after taking into consideration the opinions of the referees.

Manuscripts may be submitted in pdf/MS word format to **admin@citphy.org or susil_vcsarkar@yahoo.co.in** Online submission of the paper through our **website: www.citphy.org** is also accepted. The file should be prepared with 2.5 cm margin on all sides and a line spacing of 1.5.

The title of the paper should be short and self-explanatory. All the papers must have an abstract of not more than 200 words, the abstract page must not be a part of the main file. Abstract should be self-contained. It should be clear, concise and informative giving the scope of the research and significant results reported in the paper. Below the abstract four to six key words must be provided for indexing and information retrieval.

The main file should be divided into sections (and sub-sections, if necessary) starting preferably with introduction and ending with conclusion. Displayed formula must be clearly typed (with symbols defined) each on a separate line and well-separated from the adjacent text. Equations should be numbered with on the right-hand side consecutively throughout the text. Figures and Tables with captions should be numbered in Arabic numerals in the order of occurrence in the text and these should be embedded at appropriate places in the text. Associated symbols must invariably follow SI practice.

References should be cited in the text by the Arabic numerals as superscript. All the references to the published papers should be numbered serially by Arabic numerals and given at the end of the paper. Each reference should include the author's name, title, abbreviated name of the journal, volume number, year of publication, page numbers as in the simple citation given below :

For Periodicals : Sen, N. R. - On decay of energy spectrum of Isotopic Turbulence, 1. Appl. Phys. **28,** No. 10, 109-110 (1999).

1. Mikhilin, S. G. - Integral Equations, Pergamon Press, New York (1964).

2. Hinze, A. K. - Turbulence Study of Distributed Turbulent Boundary Layer Flow, Ph. D, Thesis, Rorke University (1970).

The corresponding author will receive page proof, typically as a pdf file. The proof should be checked carefully and returned to the editorial office within two or three days. Corrections to the proof should be restricted to printing errors and made according to standard practice. At this stage any modifications (if any) made in the text should be highlighted.

To support the cost of publication of the journal, the authors (or their Institutions) are requested to pay publication charge ` 200/- per printed page for authors of Indian Institutes and US$ 20 for others. Publication charges to be sent directly to **CALCUTTA INSTITUTE OF THEORETICAL PHYSICS, 'BIGNAN KUTIR', 4/1 MOHAN BAGAN LANE, KOLKATA-700004, INDIA.**

A pdf of the final publisher's version of the paper will be sent to the corresponding author.

All communications are to be sent to the Secretary, Calcutta Institute of Theoretical Physics, 'BignanKutir', 4/1, Mohan Bagan Lane, Kolkata-700004, India. E-mail:susil_vcsarkar@yahoo.co.in

For details please visit our website www.citphy.org

INDIAN JOURNAL OF THEORETICAL PHYSICS

BOARD OF EDITORS

CALCUTTA INSTITUTE OF THEORETICAL PHYSICS

(Formerly, Institute of Theoretical Physics)

[Established in 1953 by Late Prof. K. C. Kar, D. Sc.]

PUBLICATIONS
OF
CALCUTTA INSTITUTE OF THEORETICAL PHYSICS
"BIGNAN KUTIR"
4/1, Mohan Bagan Lane, Kolkata-700 004, India
Phone : +91-33-25555726

INDIAN JOURNAL OF THEORETICAL PHYSICS (ISSN : 0019-5693)
Research Journal containing Original Papers, Review Articles and Letters to the Editor is published quarterly in March, June, September and December and circulated all over the world.

Subscription Rates
1500 per volume (for Bonafide Indian Party)
US$ 350 (for Foreign Party)

Back Volume Rates
1500 per volume (for Bonafide Indian Party)
US$ 350 per volume or Equivalent Pounds per volume

Books Written by Prof. K. C. Kar, D. Sc.

- **INTRODUCTION TO THEORETICAL PHYSICS [Vol. I and Vol. II (Acoustics)]** Useful to students of higher physics

  Price :` 60 or US $ 10 per volume

- **WAVE STATISTICS : Its principles and Applications [Vol. I and Vol. II]** Useful to Post Graduate and Research students

  Price :` 80 or US $ 12

- **STATISTICAL MECHANICS : PRINCIPLES AND APPLICATIONS [Vol. I and Vol. II]** Useful to Advanced students of theoretical Physics

  Price :` 120 or US $ 15

- **A NEW APPROACH TO THE THEORY OF RELATIVITY [Vol. I and Vol. II]** Useful to Post Graduate and advanced students

  Price :` 50 or US $ 8

Order may be sent directly to Calcutta Institute of Theoretical Physics
"BignanKutir", 4/1, Mohan Bagan Lane, Kolkata-700 004, India