Report On

Two Days National Workshop

"Analysis of Biological Data using R- Studio"

8th& 15th April 2018



Organized by Departments of Biomedical Engineering andBiotechnology

Mahatma Gandhi Mission's College of Engineering & Technology

in association with

Biomedical Engineering Society of India (BMESI)

Chief Patron

Shri. Kamalkishor N. Kadam Chairman, Mahatma Gandhi Mission's Trust &

Former Education Minister, Govt. of Maharashtra

Patrons

Dr. Sudhir N. Kadam Vice Chancellor, MGM Institute of Health Sciences (MGMIHS)

Dr. Nitin N. Kadam Managing Trustee, MGM Trust

Dr. K. G. Narayankhedkar Director General, MGM Group of Engineering Colleges

Dr. S. K. Narayankhedkar Principal, MGM's College of Engineering & Technology

(MGMCET)

Advisory Committee

Dr. Sandhya Agarwal HOD, BME & BT Department

Dr. G.D. Jindal Professor, BME, MGMCET

Mr. U.R. Bagal Asst. Professor, BME, MGMCET

Dr. Kshipra Deshmukh Asst. Professor, BT, MGMCET

Dr. Niranjan Khambete Secretary, BMESI & Manager, Clinical Engineering,

Deenanath Mangeshkar Hospital & Research Centre

Prof. G. Muralidhar Bairy Jt. Secretary, BMESI &Asso. Prof., BME, MIT, Manipal

Organizing Committee

Prof. Jyothi Warrier Asst. Prof., BME, MGMCET & Convener, NWABDR

Prof. Sandeep Singh Asst. Prof., BT, MGMCET & Co-Convener, NWABDR

Prof. Nishant Patil Asst. Prof., BME, MGMCET & Secretary, NWABDR

Prof. Priyanka Jadhwar Asst. Prof., BME, MGMCET & Jt. Secretary, NWABDR

Prof. Parvathy Krishna Asst. Prof., BT, MGMCET & Treasurer, NWABDR

Prof. Preeti Athavale Asst. Prof., BME, MGMCET & Member, NWABDR

Prof. Shruti Bajpai Asst. Prof., BT, MGMCET & Member, NWABDR

Prof. Shweta Choudhari Asst. Prof., BT, MGMCET & Member, NWABDR

Prof. Akanksha Rajpoot Asst. Prof., BT, MGMCET & Member, NWABDR

Prof. Ashwini Tidke Asst. Prof., BME, MGMCET & Member, NWABDR

Prof. Shashikant Rathod Asst. Prof., BME, MGMCET & Member, NWABDR

Prof. Arbacheena Bora Asst. Prof., BME, MGMCET & Member, NWABDR

Preface

Medicine and biomedical sciences have become data-intensive fields, which, at the same time, enable the application of data-driven approaches and require sophisticated data analysis methods. Biomedical informatics provides a proper interdisciplinary context to integrate data and knowledge when processing available information, with the aim of giving effective decision-making support in clinics and translational research. Statistics and statistical analysis plays a crucial role in research and analyzing of biological data.R is a programming language that was developed some decades ago by a couple of statisticians in New Zealand. Because it was developed by statisticians, it is very suitable for statistical programming. Together with R-Studio it makes a fantastic environment for data-related programming. Even if one is not familiar with programming at all, we can think of a programming language, such as R, as something that lives in a computer, and when requested, performs computation on the input given to it. In order to give input to R, we need to open a console that allows typing inputs to R. The same console also prints the outputs of the computations.

R-Studio is a free and open-source integrated development environment (IDE) for R, a programming language for statistical computing and graphics. R-Studio was founded by JJ Allaire, creator of the programming language . R-Studio is available in open source and commercial editions and runs on the desktop (Windows, macOS, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian, Ubuntu, Red Hat Linux, CentOS, openSUSE and SLES).R-Studio is written in the C++ programming language and uses the framework for its graphical user interface. Work on R-Studio started around December 2010 and the first public beta version (v0.92) was officially announced in February 2011. Version 1.0 was released on 1 November 2016 and Version 1.1 was released on 9 October 2017.

Acknowledgement

We take this opportunity to express gratitude to Shri. Kamalkishor N. Kadam, Chairman, Mahatma Gandhi Mission Trust, Dr. Sudhir N. Kadam, Vice Chancellor, MGM Institute of Health Sciences (MGMIHS), Dr. Nitin N. Kadam, Managing Trustee, MGM Trust, Dr. K. G. Narayankhedkar, Director General, MGM Group of Engineering Colleges and Dr. S. K. Narayankhedkar, Principal, MGM's College of Engineering & Technology (MGMCET)who spared their valuable time to guide the deliberations of the Workshop.

We also extend our thankstoBMESI and Dr. Niranjan Khambete (Secretary, BMESI) &Prof. G. Muralidhar Bairy (Jt. Secretary, BMESI) have agreed as joint organizer for workshop and also providing sponsorship for the same.

On behalf of MGM's CET, and the Departments of Biomedical Engineering & Biotechnology, we extend a very hearty thanks to the expert resource person Dr. Amiya Bhowmickfor sharing with us his experience, findings and opinions on R-studio.

We also express thanks to Dr. Sandhya Agarwal, HoD, Departments of BME & BT for her enormous cooperation in the organization of this event. We would also like to thank Dr. Sankar, HoD, Computer Engineering Department for providing us the technical support as and when required.

Organising a national workshop which includes many activities, in itself is a difficult task. This difficult task had been made easy by my colleagues and all the members of the Advisory and Organising committee. We would like to thank all committee members for their support and help.

With a great sense of pleasure and satisfaction, we express sincere thanks to the participants of the National Workshop, without their participation and help; the organisation of national workshop would have not been possible.

Contents

S. No	Topic	Page No.
1	Programme Schedule	1
2	Executive Summary	3
3	Introduction	4
4	Objectives	5
5	Detailed Report	6
6	Conclusion	11
7	List of Participants	12
8	List of Guest Speaker and his CV	14
9	Feedback of the Participants	21
10	Receipts and Expenditure Statement	22
11	Appendix I	25

1. Programme Schedule

Day 1 (April 8, Sunday)

Day 1 (April 8, Sun Time	Subject		
DAY-I	The state of the s		
9:00AM-10:00AM	Registration		
DAY-I	T (IEP'')		
10:00AM-10:30AM	Inauguration and Felicitation		
10:30 AM – 11:00 AM	Tea Break		
	Introduction to R and R-Studio		
	 Installation and different options in RStudio, customizing the environment 		
DAY – I (Session – A)	 Installing Packages (install.packages), loading packages (require, 		
11:00 AM – 12:00 PM	library), Looking into help files		
(ARB)	■ Familiar with options (setwd, getwd, sessionInfo)		
, ,	■ Data import and export (read.table, read.csv, write.table,		
	write.csv)		
	Handling data in R		
	Basic data types: numeric, integer, logical, character		
	 Data storage facilities in R (c(), vector, matrix, data.frame, 		
	list, rep, numeric, seq)		
DAY – I (Session – B)	Difference between matrix and data.frame		
12:00 PM – 01:30 PM	Subsetting and modifying data (select, subset, filter, which,		
(ARB)	reshape)		
	Combining data (rbind, cbind, merge)		
	 Some useful functions (is.na, dim, complete.cases, summary, 		
	aggregate, class)		
01:30 PM - 02:30 PM	Lunch Break		
01.301141 02.301141	 Visualizing data distributions: Histogram, boxplot (hist, boxplot, plot) 		
1	Probability mass functions and Probability density functions		
	Probability mass functions and Probability density functions Area under the curve (Integrate)		
	Area under the curve (Integrate)		
DAY – I (Session – C)	Area under the curve (Integrate)Binomial distribution, Uniform distribution, Normal distribution, t- distribution		
DAY – I (Session – C) 02:30 PM – 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) 		
02:30 PM - 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, 		
	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, 		
02:30 PM - 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) 		
02:30 PM - 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions 		
02:30 PM - 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions 		
02:30 PM – 04:00 PM (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych 		
02:30 PM - 04:00 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break		
02:30 PM – 04:00 PM (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) 		
02:30 PM – 04:00 PM (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) 		
02:30 PM – 04:00 PM (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) 		
02:30 PM – 04:00 PM (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) Correlation (cor, corr.test) 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM DAY - I (Session - D)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) Correlation (cor, corr.test) Checking model assumptions and associated plots (boxcox) 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) Correlation (cor, corr.test) Checking model assumptions and associated plots (boxcox) Regression diagnostics and associated plots (influential points, high-leverage 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM DAY - I (Session - D) (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) Correlation (cor, corr.test) Checking model assumptions and associated plots (boxcox) Regression diagnostics and associated plots (influential points, high-leverage points, non-constant variance or error, correlation of errors, possible nonlinearity) 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM DAY - I (Session - D) (ARB)	 Area under the curve (Integrate) Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) Strong emphasis on visualizing distribution functions Important packages: Hmisc, psych Tea Break Model building in R - I(Simple Linear Regression) Train and test partition of data (sample, sample.split) Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) Correlation (cor, corr.test) Checking model assumptions and associated plots (boxcox) Regression diagnostics and associated plots (influential points, high-leverage points, non-constant variance or error, correlation of errors, possible nonlinearity) 		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM DAY - I (Session - D) (ARB)	■ Area under the curve (Integrate) ■ Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) ■ Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) ■ Strong emphasis on visualizing distribution functions ■ Important packages: Hmisc, psych ■ Tea Break Model building in R - I(Simple Linear Regression) ■ Train and test partition of data (sample, sample.split) ■ Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) ■ Correlation (cor, corr.test) ■ Checking model assumptions and associated plots (boxcox) ■ Regression diagnostics and associated plots (influential points, high-leverage points, non-constant variance or error, correlation of errors, possible nonlinearity) ■ Prediction using model for new data		
02:30 PM - 04:00 PM (ARB) 04:00 PM - 04:30 PM DAY - I (Session - D) (ARB)	■ Area under the curve (Integrate) ■ Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm) ■ Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var) ■ Strong emphasis on visualizing distribution functions ■ Important packages: Hmisc, psych ■ Tea Break Model building in R – I(Simple Linear Regression) ■ Train and test partition of data (sample, sample.split) ■ Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary) ■ Correlation (cor, corr.test) ■ Checking model assumptions and associated plots (boxcox) ■ Regression diagnostics and associated plots (influential points, high-leverage points, non-constant variance or error, correlation of errors, possible nonlinearity) ■ Prediction using model for new data ■ Measures of Model accuracy (R², Adjusted R², AIC, BIC, Mallow's Cp)		

Day 2 (April 15, Sunday)

Day 2 (April 15, Sunday)			
Time	Topics		
DAY – II (Session – A) 09:30 AM – 11:00 AM (ARB)	 Model building in R - II(Multiple Linear Regression) Model building with more than one explanatory variable Variable selection, exploring correlations (corplot, pairs) Comparing models (anova) Forward, Backward and Mixed selection (regsubsets), selection criteria Prediction for new data using the best model (predict, predict.lm), confidence interval and prediction interval Regression diagnostics and checking model assumptions (Multicollinearity, Variance Inflation Factor, vif) (connect - Session IID) Important Packages: leaps, caret, corrplot, ISLR, MASS Useful functions: createDataPartition, stepAIC 		
11:00 AM – 11:30 AM	■ Tea Break		
DAY – II (Session – B) 11:30 AM – 01:30 AM (ARB)	 Introduction to basic R programming The basic syntax for programming in R (function) Create basic functions Utility of writing functions (Toy examples for biologists) Loops and control flow statements and basic examples (for, if) Application to Bootstrapping 		
01:00 PM - 02:30 PM	■ Lunch Break		
DAY – II (Session – C) 02:30 AM – 04:00 AM (ARB)	Some Resampling methods in Biology (Model selection) Validation set approach Cross validation (Loocy, K-fold) Bootstrap, Jacknife (boot) Bootstrapping regression model (alr3, alr4)		
04:00 PM – 04:30 PM	■ Tea Break		
DAY – II (Session – D) 04:30 PM – 06:00 PM (ARB)	 Discussion of some Case studies from Biology Valedictory 		

2. Executive Summary

MGM's College of Engineering &Technology was established in 1986 by Hon. Shri Kamal Kishore Kadam (M.Tech., IIT-Mumbai, Ex-Minister of Education, Maharashtra) with a futuristic vision to provide qualitative education by applying innovative and dynamic techniques. This ISO 9001-2000 certified college is affiliated to University of Mumbai and approved by AICTE, New Delhi.

Bio-Medical Engineering department has been running Undergraduate courses since 1986. It has been often getting RANK in University of Mumbai. Bio-Medical Engineering department has been running Postgraduate courses and PhD program since 2000 and 2016 respectively. Sixty PG projects in 24 research area have led to good number of publications .The Department of Biotechnology was established in 2007. The department is one of its kinds imparting quality education and promoting goal-orientated interdisciplinary research by interfacing modern biology with upcoming technologies. The program is designed to train students to deal with any technological applications that use biological systems. Graduate qualification in biotechnology can lead to placements in research laboratories run by government and corporate sector.

Biomedical Engineering society of India (BMESI) is a non-profit professional All India Organization having headquarters at Department of Biomedical Engineering, Manipal Institute of Technology, Manipal. It organizes symposia, conferences, lectures, study programs etc. covering broad areas of biomedical engineering. It establishes contacts with other learned and professional organizations, and publishes newsletters, official journals and magazines. Expert advisory panels are set and collaboration is made with pool of eminent researchers and scholars in the field of biomedical engineering throughout the country. It continuously strives to improve the standards, terminology, equipment, methods and safety practices for state of the art health care to the people.

National Workshop on "Analysis of Biological Data Using R-Studio" was organized by Departments of Bio-Medical Engineering (BME) & Bio-Technology (BT), MGMCET in association with Bio-Medical Engineering Society of India (BMESI) on 08th APRIL 2018 & 15th APRIL 2018. The workshop was formally inaugurated by Dr. K.G. Narayankhedkar, Hon'ble Director General, MGMCET, Kamothe, Navi Mumbai. The inaugural function was presided by Dr. S.K. Narayankhedkar, Principal, MGMCET, Kamothe, Navi Mumbai. There were 52delegates participated in the workshop and learnt about Introduction to R and R-Studio, Handling data in R, Model building in R – I (Simple Linear Regression), Model building in R – II (Multiple Linear Regression), Statistical analysis using R-studio, Introduction to basic R programming, Some Re-sampling methods in Biology (Model selection), Discussion of some Case studies from Biology. The Expert Resource Person for the workshop was Dr. Amiya Bhowmick (Institute of Chemical Technology, Mumbai, India) and Ms Jyoti V Jethe, research scholar from MGMCET, assisted the resource person in hands-on sessions.

3. Introduction

R-Studio is a free and open-source integrated development environment (IDE) for R, a programming language for statistical computing and graphics. The significance of Biological data and its effective analysis is unavoidable to have a good piece of research and to make a fruitful decision. From students to scholars, individual researchers to research institutes throughout the world come across this data analysis part to carry out their research. Statistics and statistical analysis plays a crucial role in research and analyzing biological data.R-Studio is a full-blown integrated development environment (IDE), and it is open source software. The Scope of workshop is to cover theoretical discussion and practical exposure on the following aspects of statistical analysis using R: Introduction to R & R-Studio; Data types in R; Importing and handling data in R; Working with functions in R; Exploratory Data Analysis using R (Data Visualization); Descriptive Statistics (summary Measures) using R; Simple and Multiple Linear Regression Analysis; Resampling methods in Biology (Model selection); Discussion of some Case studies from Biology.

4. Objective

The aim of this workshop is to provide hands on training in understanding the concepts in data science and statistical analysis of biological data using R programming which is one of the most widely used open source package in the field of Data Science across the world. This workshop will help the participants in applying proper research methodology and improving their analysis by applying various statistical techniques using R programming.

5. Detailed Report

The workshop covered various topics Installation and different options in R-Studio, customizing the environment, Installing Packages (install.packages), loading packages (require, library), Looking into help files, Data import and export (read.table, read.csv, write.table, write.csv, Basic data types: numeric, integer, logical, character, Data storage facilities in R (c(), vector, matrix, data.frame, list, rep, numeric, seq), Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (lm, summary), etc.. The participants wereprovided hands-on practice on the topic mentioned.Training Modules included are

- Introduction to R and R-Studio
- Handling data in R
- Model building in R I (Simple Linear Regression)
- Model building in R II (Multiple Linear Regression)
- Introduction to basic R programming
- Some Re-sampling methods in Biology (Model selection)
- Discussion of some Case studies from Biology

R is one of the most widely used platforms in research. R is the name of the programming language, and the software it runs in. There are many reasons to learn and use R

- The software is free and the algorithms used in base R are well documented and trusted
- Any additional packages published on CRAN are also well documented, so that algorithms can be verified.
- The process of working in R relies on writing code. Using code rather than dropdown menus means that you have a record of the work that you've done (to re-visit later for edits or to remind you of what you did). You can also include comments in your code to remind yourself or explain to others what you have done. Code can be easily emailed or shared among collaborators.
- Extensive help is available for R. Help is included in base R and in packages. Numerous tutorials are available online. Several user groups exist online that allow people to ask and answer questions. Many solutions to problems can be gleaned from the pages documenting these user group questions.
- R is highly extensible through packages created by R users and published online.
 These packages are usually highly specific, providing tools fit for purpose. Most
 packages are well documented, including their methods and algorithms, and people
 can thus verify their accuracy and place their trust in those packages. However, this is
 not a given.
- R is excellent for producing highly customisable figures that can be saved in a variety of formats.
- One can conduct almost all of one's project workflow entirely in R, from data entry and management, to analysis, and even to producing presentations, posters, manuscripts, and books.
- A script will be generated (so you can re-run your past analyses); which also means
 the script can be used as a template for modification into a new script (viz. code
 recycling).

- Code checking where syntax-errors and non-existent functions/variables/data frames will be identified in existing scripts.
- Code-completion for new commands (or when editing/re-writing existing commands).
- Libraries (collection of commands; add-ins as they are sometimes called) already installed can be loaded (and unloaded) without writing commands.
- Libraries that need to be downloaded over the internet can be downloaded-andunpacked without writing commands (and all extra required-libraries are automatically downloaded).
- A history-loge exists so prior-run commands can easily be found and re-run.
- All graphs are kept for later export (R only keeps the last graph, and overwrites it whenever a new graph is produced).
- You can have multiple scripts open at a time, so you can copy different lines of codes from different scripts in order to create your new script (code recycling).

The following practical topics were covered during two days workshop

Sr. Topic

Introduction to R and RStudio

- Installation and different options in RStudio, customizing the environment
- Installing Packages (install.packages), loading packages (require, library),
 Looking into help files
- Familiar with options (setwd, getwd, sessionInfo)
- Data import and export (read.table, read.csv, write.table, write.csv)

Handling data in R

1

3

4

- Basic data types: numeric, integer, logical, character
- Data storage facilities in R (c(), vector, matrix, data.frame, list, rep, numeric, seq)
- 2 Difference between matrix and data.frame
 - Subsetting and modifying data (select, subset, filter, which, reshape)
 - Combining data (rbind, cbind, merge)
 - Some useful functions (is.na, dim, complete.cases, summary, aggregate, class)
 - Visualizing data distributions: Histogram, boxplot (hist, boxplot, plot)
 - Probability mass functions and Probability density functions
 - Area under the curve (Integrate)
 - Binomial distribution, Uniform distribution, Normal distribution, t- distribution (rbinom, runif, rnorm, dnorm, qnorm, pnorm)
 - Skewed distribution, descriptive measures (mean, median, mode, quantile, range, sd, kurtosis, skew, describe, min, max, var)
 - Strong emphasis on visualizing distribution functions
 - Important packages: Hmisc, psych

Model building in R – I (Simple Linear Regression)

- Train and test partition of data (sample, sample.split)
- Simple Linear Regression (Concepts, Assumptions, Understanding R outputs) (1m, summary)
- Correlation(cor, corr.test)
- Checking model assumptions and associated plots (boxcox)
- Regression diagnostics and associated plots (influential points, high-leverage points, nonconstant variance or error, correlation of errors, possible nonlinearity)

- Prediction using model for new data
- Measures of Model accuracy (R², Adjusted R², AIC, BIC, Mallow's Cp)
- Important package: MASS, car
- Important functions: outlierTest, qqPlot, leveragePlots, influencePlot, ncvTest

Model building in R – II (Multiple Linear Regression)

- Model building with more than one explanatory variable
- Variable selection, exploring correlations (corplot, pairs)
- Comparing models (anova)
- Forward, Backward and Mixed selection (regsubsets), selection criteria
- 5 Prediction for new data using the best model (predict, predict.lm), confidence interval and prediction interval
 - Regression diagnostics and checking model assumptions (Multicollinearity, Variance Inflation Factor, vif) (connect - Session IID)
 - Important Packages: leaps, caret, corrplot, ISLR, MASS
 - Useful functions: createDataPartition, stepAIC

Introduction to basic R programming

- The basic syntax for programming in R (function)
- Create basic functions
- Utility of writing functions (Toy examples for biologists)
 - Loops and control flow statements and basic examples (for, if)
 - Application to Bootstrapping

Some Resampling methods in Biology (Model selection)

- Validation set approach
- 7 Cross validation (Loocy, K-fold)
 - Bootstrap, Jacknife (boot)
 - Bootstrapping regression model (alr3, alr4)
- 8 Discussion of some Case studies from Biology

PHOTOS:





















6. Conclusion

At the end of the workshop, participants got familiar with R programming language; Understood statistical analysis using R programming; Learnt how to explore research on biological data using R programming and how to represent biological data using basic graphs in R.

Participants also learnt different techniques of Handling data in R, Model building in R-I (Simple Linear Regression), Model building in R-II (Multiple Linear Regression), Introduction to basic R programming, Some Re-sampling methods in Biology (Model selection), Discussion of some Case studies from Biology.

The workshop has been very useful for all the participants as they were confident of using these tools in their career.

7. List of Participants

Sr. No	Name	College/Industry
1.	Abhisekh Mankar	MGMCET
2.	Aeshaben Gandhi	MGMCET
3.	Afifa Juhi Khan	MGMCET
4.	Akhidev Devdas	MGMCET
5.	Annesha Moza	MGMCET
6.	Anisha Jagdale	MGMCET
7.	Aparna Lakhe	MGMCET
8.	Arbacheena Bora	MGMCET
9.	Ashraf Firoz Shaikh	MGMCET
10.	Ashwini Tidke	MGMCET
11.	Chetan Ippar	MGMCET
12.	Claire Vania	MGMCET
13.	Deepa Kale	MGMCET
14.	Ganesh Teke	MGMCET
15.	Gomati Nandekar	MGMCET
16.	Harshal Raghotre	MGMCET
17.	Jjismi Easow	MGMCET
18.	Jyoti Jethe	MGMCET
19.	Kanan Dombgare	MGMCET
20.	Khushboo Dhuri	MGMCET
21.	Krutika Jadhav	MGMCET
22.	Krutika Jagdale	MGMCET
23.	Manasi Sawant	MGMCET
24.	Mansi Gaikwad	MGMCET
25.	Mayur Ghangale	TATA ELXSI PVT. LTD, PUNE
26.	Mayuri Mulay	MGMCET
27.	Mohini Mehta	MGMCET
28.	Nikita Panchal	VEOL MEDICAL TECHNOLOGIES PVT. LTD
29.	Nikita Panyam	MGMCET
30.	Nishant Patil	MGMCET
31.	Nischita Moundekar	MGMCET
32.	Parth Gamare	MGMCET
33.	Prajakta Mhatre	MGMCET
34.	Priya Gaikar	MGMCET
35.	Pumpy Kathar	MGMCET
36.	Rajpal Jagtap	MGMCET
37.	Salik Khan	MGMCET
38.	Shashikant Rathod	MGMCET
39.	Sharda Tondare	BVU COLLEGE OF ENGINEERING, PUNE
40.	Shrikant Patil	MGMCET
41.	Shruthy Seshadrinathan	MGMCET
42.	Shruti Bajpai	MGMCET
43.	Shruti Gadkari	MGMCET
44.	Shweta Chaudhari	MGMCET

45.	Shweta Singh	MGMCET
46.	Siddhant Ojha	MGMCET
47	Simanjay Singh	SELF EMPLOYED
48.	Sneha Bansod	MGMCET
49.	Sushma Bhat	BARC,MUMBAI
50.	Sushma Nair	MGMCET
51.	Tejasvini Shendge	MGMCET
52.	Uday Shirsekar	MGMCET

8. List of Guest Speakers

Dr. Amiya R. Bhowmick Asst. Professor, Mathematics Dept, ICT, Mumbai.

Amiya Ranjan Bhowmick, PhD

Webpage : https://sites.google.com/site/amiyaiitb/

Email :amiyaiitb@gmail.com

Mobile :+91 -7738101583/8334835300

Date of Birth :12th February, 1985

Gender :Male
Marital Status :Married
Nationality :Indian



Research Interest

My research interest includes application of growth models in biological research, empirical assessment of population dynamics with potential application in ecology and conservation biology. Some specific areas include: 1) Growth curve models, 2) Ecological Niche Modelling, 3) Allee effect 4) Stochastic Population Dynamics.

PhD Thesis

Title: An Extended Family of Density Dependent Growth Curve Models and Related Ecological Issues **Advisor:** Dr. Sabyasachi Bhattacharya, Associate Professor, Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata

Current Employment (March 16, 2015 - Present)

Institute of Chemical Technology, Mumbai, Nathalal Parekh Marg, Matunga - 400019

Assistant Professor, Department of Mathematics

Courses Taught/Teaching (More details in Personal Webpage):

1.Postgraduate Level

- a) Applied Statistics-I (Introduction to Probability, Theory of Point Estimation)
- b) Applied Statistics II (Statistical Inference, Regression Analysis, Statistical Simulation),
- c) Applied Statistics III (Time Series and Statistical Quality Control),
- d)Bioinformatics and Statistics Lab (https://sites.google.com/site/amiyaiitb/teaching/statistics-lab)
- e)Mathematical Biology (https://sites.google.com/site/amiyaiitb/teaching/mathematical-biology)
- f)Advanced Real Analysis (Measure Theory and Integration)

2. Undergraduate Level (B. Tech and B.Chem Engg.)

- a) Applied Mathematics I
- b)Applied Mathematics II
- c)Engineering Application of Computer.

Project Supervision:

1.M.Sc. Home paper project

a)Study of some resampling techniques and their application in Artificial Neural Network using R (Steffi D'Souza, 2016)

b)Ecological Niche Modeling using Logistic, Lasso and Ridge Logistic Regression: A case study using *Mikania micrantha* kunth (Shobhana Gopal Iyer, 2017)

c)Prediction of Animal Movement using Hidden Markov Model (Sunil Gupta, 2017).

d)Study of growth curve models with continuously varying parameters with application to real data (Amir Razza Khan, 2018, ongoing)

e)Study of growth curve models with randomly varying parameters with application to real data (Harish Nagula, 2018, ongoing)

Previous Employment

- 1.Indian Statistical Institute, Kolkata (March 2010 March 2015) Research Fellow at Agricultural & Ecological Research Unit.
- 2. Cytel Statistical Software Services Pvt. Ltd. (2008-2009) Statistical Programmer

Journal Reviewer

- 1. Ecological Modelling, Elsevier publications.
- 2. Mathematical Biosciences, Elsevier publications.
- 3. Stochastic Environmental Research and Risk Assessment, Springer publications.

Journal Publications

- 1.Pal, A., **Bhowmick, A. R.**, Yeasmin, F, and Bhattacharya, S. Evolution of Model Specific Relative Growth Rate: Its Genesis and Performance Over Fisher's Growth Rates. Journal of Theoretical Biology (2018) (doi: 10.1016/j.jtbi.2018.02.012) Vol-444, pp 11-27 (I.F. 2.260).
- 2. Chakraborty, B., Bhowmick, A. R., Chattopadhyay, J. and Bhattacharya, S. Physiological Responses of Fish under Environmental stress and Extension of Growth (curve) models. Accepted (2017) in Ecological Modelling (Elsevier: Impact Factor: 2.33, SCI).
- 3.Mukhopadhyay, S., Hazra, A, **Bhowmick, A. R.** and Bhattacharya, S. On comparison of Relative growth rates under different environmental conditions with application to biological data. Metron (2016) (Springer) Volume 74, Issue 3, pp 311–337
- 4.Chattopadhyay, A., Sahah, B., **Bhowmick, A. R.** and Bhattacharya, S. Allee effect and associated risk of species extinction: An empirical study based on global population dynamics database. Nonlinear Studies 23(1) 1-15 (2016)
- 5.**Bhowmick, A. R.,** Bandopadhyay, S., Rana, S. And Bhattacharya, S. A Simple Approximation of Moments of the Quasi-equilibrium Distribution of an Extended Stochastic theta-Logistic Model with Non-integer Powers Mathematical Biosciences 271, 96-112 (2016) (Elsevier: Impact Factor: 1.454, SCI) (Part of Thesis)
- 6.**Bhowmick, A. R.**, Saha, B., Chattopadhyay, J., Ray, S. and Bhattacharya, S. Cooperation in Species: Interplay of Population Regulation and Extinction through Global Population Dynamics Database. Ecological Modelling 312, 150-165 (2015) doi-10.1016/j.ecolmodel.2015.05.023) (Elsevier: Impact Factor: 2.33, SCI) (Part of thesis).
- 7.Sasmal, K. S., **Bhowmick, A. R.,** Bhattacharya, S and Chattopadhyay, J. Interplay of functional responses and weak Allee effect on pest control via viral infection or natural predator An eco-epidemiological study. Differential Equation and Dynamical Systems 24(1),(2015) 21-50(Springer:doi-10.1007/s12591-015-0240-3)
- 8.Kang, Y., Sasmal, S. K., **Bhowmick, A. R.** and Chattopadhyay. J. A Host-Parasitoid System with Predation-Driven Component Allee effects in Host Population, Journal of Biological Dynamics 9, 213-232 (2014)(Impact factor: 1.147, Taylor and Francis).
- 9.Rana, S., **Bhowmick**, **A. R.** and Bhattacharya, S. Impact of Prey Refuge on a Discrete Time Predator-Prey System with Allee Effect. International Journal of Bifurcation and Chaos (2014, 24(9) (World Scientific Publishing, Impact Factor: 0.921, SCI)
- 10.**Bhowmick, A. R.,** Bhattacharya, S. A New Growth Curve Model for Biological Growth: Some Inferential Studies on the Growth of *Cirrhinus mrigala*. Mathematical Biosciences 254 (28-41) (2014) DOI: 10.1016/j.mbs.2014.06.004 (Elsevier: Impact Factor: 1.454, SCI) (Part of Thesis)
- 11.**Bhowmick, A. R.,** Saha, B., Chattopadhyaya, J. and Bhattacharya, S. A Modified Theta-logistic Model with Cooperation for Understanding Species Extinction. Journal of the Indian Society of Agricultural Statistics (Special Issue 2014) 68(2) 39-54. (Part of Thesis)

- 12.**Bhowmick, A. R.**, Chattopadhyaya, G. and Bhattacharya, S. Simultaneous Identification of Growth Law and Estimation of Its Rate Parameter: A New Metric. Journal of Biological Physics 2014. 40(1) 71-95. (Springer: Impact Factor: 0.951, SCI) (Part of Thesis)
- 13. Sardar, T., Mukhopadhyay, S., **Bhowmick, A. R.** and Chattopadhyay, J. An Optimal Cost Effective Study on Zimbabwe Cholera Seasonal Data from 2008-2011. PLOS ONE. (2013) 8(12) (Impact Factor: 3.73)
- 14.Kang, Y., Sasmal, S., **Bhowmick, A. R.** and Chattopadhyay, J. Dynamics of a Predator-Prey System with Prey Subject to Allee effect and Disease. Mathematical Biosciences and Engineering. Pages: 877 918, Volume 11, Issue 4, 2014 doi:10.3934/mbe.2014.11.877(AIMS: Impact Factor:1.195)
- 15. Saha, B. and **Bhowmick**, A. R. Predicting the Extinction Vulnerability of Species: Stochastic Approach to Deterministic Population Dynamic Models. Bulletin of Calcutta Mathematical Society. (2013) Vol 105 (3).
- 16.Saha, B., **Bhowmick, A. R.**, Chattopadhyay, J. and Bhattacharya, S. On the evidence of an Allee effect in Herring Populations and consequences for population survival: A model based study. Ecological Modelling. Volume 250. Pages72-80. (2013) (Elsevier: Impact Factor: 2.33, SCI)

Conference Proceedings

1.Saha, B., **Bhowmick, A. R.,** Chattopadhyaya, J. and Bhattacharya, S. Correlated noises in dynamical system with θ -logistic growth process of species. International Conference on Advances in Pure and Applied Mathematics "ICAPAM-2014". International Journal of Technology. **ISSN 2086-9614.**

Magazine Writing

1.Evidence of Cooperation among Animal Taxa: A Protective Mechanism against Species Extinction: An Empirical Study based on Global Population Dynamics Database. **Green Earth** published by Millenium Institute of Technology, Kolkata, India.

Article under review/submitted

- 1.Climatic Indicators for Predicting Species Invasion: An Empirical Investigation of the Spread of \emph{Mikania micrantha} Kunth in India (with Banerjee, A. and Iyer, S. G.)
- 2.Environmental factors as indicators of dissolved oxygen concentration and zooplankton abundance: Deep learning versus traditional regression approach (with Arnab Bajerjee, Santanu Ray)

Article under preparation

- 1.A machine learning approach for modelling physiological response to fish (with Mukherjee, J.)
- 2.Understanding Growth Regulation of Natural Populations by Extended Family of Growth Models using Fractional Derivatives: Cases Studies from the Global Population Dynamics Database (with T. Sardar and Bhattacharya, S.)

Academic Visit

- $1.(26 \text{th May}\ 2016-4 \text{th June}\ 2016)$ Academic visit to Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata.
- 2.(6th March 2017 07th March 201) Academic visit to Department of Ecology and Conservation Biology, Periyar Tiger Reserve, Thekkady, Kerala.
- $3.(23_{rd}\,March\ 2017-25_{th}\,March\ 2017)$ Academic visit to the Department of Mathematics, BITS Pilani, Goa campus.

Workshops and Conferences

Speaker/Resource Person/Coordinator

- 1.Acted as a core organizing committee member and a resource person for the Workshop and Statistical Methods and R Programming for Biologists during $07_{th}-13_{th}$ March 2018 organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata.
- 2. Invited to deliver a series of lectures on R Programming for Biologists at the Rajib Gandhi University,
- Arunachal Pradesh during the North East Workshop on Modern Ecological and Agricultural Practices with Statistical Methodologies and R Software-2018 (26th -27th February 2018).

- 3.Acted as a resource person for the 'Dr Ganesh Prasad Math Fest 2018' organised on Tuesday, 23rd Jan 2018 as part of 'Indian Mathematician Chair' instituted in the Dept of Maths & Stats, Guru Nanak Khalsa College of Arts, Science and Commerce. Delivered a keynote address on "Applications of Mathematics".
- 4.Resource person on R Programming in the Short Term Training Programme on "Applicable Mathematics and Statistical Techniques" organized by the Basic Sciences and Humanities Department, Don Bosco Institute of Technology, Mumbai during 2nd 6nd January 2018.
- 5.Resource person on R Programming in the Workshop "Modern Ecological and Agricultural Practices with Statistical Methodologies and R software" held in North Eastern Hill University, Tura Campus, Meghalaya, India during 30th-31th March 2017.
- 6.Resource person on R Programming in the Workshop "Species Distribution Modelling with MaxEnt" during 9th 14th January 2017 organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata (wixsite.com/sdmworkshop).
- 7.Delivered a lecture entitled "Statistical treatment of nonlinear regression models" on 10th November 2016, in the Refresher Course organized by the Department of Statistics, University of Mumbai.
- 8. National Conference "4th India Biodiversity Meet 2015" (24 -27th October 2016) organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata jointly with Government College of Engineering and Textile Technology (*Invited Speaker*).
- 9. Coordinator of Workshop on Machine Learning Using R (Sept 30 Oct 1, 2016) organized by the Department of Mathematics, Institute of Chemical Technology, Mumbai. The workshop is financially supported by TEQIP-II
- 10. Coordinator of National Seminar on Computational and Mathematical Biology (10th 11th September, 2016) organized by the Department of Mathematics, Institute of Chemical Technology, Mumbai. The seminar series is financially supported by TEQIP-II (speaker).
- 11.Resource Person at the "Workshop on R with Applications to Quantitative Finance" organized by the Department of Mathematics, BIT Pilani, Goa with financial support from TEQIP-II (26th-28th August 2016)
- 12. Coordinator for Industry-Academic Interaction Program (14th May 2016) in the department of mathematics, ICT Mumbai (Funded by TEQIP-II) (*Organizer*).
- 13.Coordinator for SPSS Training program for M.Sc. and Ph.D. students (11th 12th May 2016) organized by the department of mathematics, ICT Mumbai jointly with Veracitiz Pvt. Ltd (Funded by TEQIP-II) (*Organizer*).
- 14.3rd National Workshop on Statistical Methods and R Programming (10-15th February 2016), jointly organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata and Department of Mathematics, Institute of Chemical Technology, Mumbai. (*Resource person and Local Coordinator at ICT, Mumbai* https://sites.google.com/site/rworkshopisi/)
- 15.Statistical Modeling and Data Analysis: Its Engineering Applications, TEQIP-II sponsored one week faculty development program (11-15th January 2016) organized by Government College of Engineering and Textile Technology, Berhampore, West Bengal (*Resource Person*)
- 16.The 2015 NNMCB National Meeting, (27-30th December 2015) National Network for Mathematical and Computational Biology, Pune Node jointly organized by NCL Pune and IISER Pune (*Invited Speaker*).
- 17.UGC Refresher Course for college teachers, University of Mumbai (5th December 2015). *Resource person for R Programming*.
- 18.International Conference "3rd India Biodiversity Meet 2015" (16 -18th November 2015) organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata jointly with Biomathematical Society of India (*Invited Speaker and member of organizing committee*).
- 19.National Workshop on Statistical Methods and R Programming (25th-29th May 2015) organized by Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata (*Speaker and member of the organizing committee*, https://sites.google.com/site/amiyaiitb/workshop)
- 20. Workshop on R-Programming for students and Researchers (16th-17th September 2014) organized by Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata (*Speaker and member of the organizing committee*).
- 21.Data Analysis and Statistics with R, 2012 organized by School of Mobile Computing, Jadavpur University in collaboration with Indian GNU/Linux Users Groups (*Resource Person*).
- 22.International Conference "India Biodiversity Meet 2014" (21st-23rd November) organized by Indian Statistical Institute, Kolkata jointly with Biomathematical Society of India. (Title: An Extended Family of

Density Dependent Growth Curve Models and Related Ecological Issue, https://sites.google.com/site/indiabiodiversitymeet2014/home) (Invited Speaker)

Workshops and Conference Attended

1.International Conference on New paradigms in Statistics for Scientific and Industrial Research Jointly Organized

by Indian Association of Productivity, Quality & Reliability and CSIR – Central Glass & Ceramic Research Institute (4th – 6th January 2018) (Invited Talk: Evolution of Model Specific Growth Rate: Theoretical Analysis and Application to Real Data)

- 2.One Week Certificate Programme on Advanced Data Analysis in Management (ADAM) (14-18th December 2015) Department of Management Studies, Indian Institute of Technology, Delhi.
- 3.15th International Symposium on Mathematical and Computational Biology (2-6th November 2015), Indian Institute of Technology, Roorkee. (*Poster presented*)
- 4.DST-SERB Summer School on Matrix Methods and Fractional Calculus, Centre for mathematical and statistical sciences. 28th April-23rd May, 2014 (Best student award for highest marks).
- 5.3rd International Symposium on Complex Dynamical. Systems and Applications (CDSA -2014). March 10 12, 2014 organized by Agricultural and Ecological Research Unit & Physics and Applied Mathematics Unit Indian Statistical Institute, Kolkata.
- 6.National Conference on Mathematical and Theoretical Biology (NCMTB 2014) (20-21st February 2014) organized by Centre for Mathematical Biology and Ecology, Department of Mathematics, Jadavpur University, Kolkata and Bio mathematical Society of India. (*Paper presented*).
- 7.Bayes by the Bay A Pedagogical workshop on Bayesian Methods in Science (4th-8th January 2013) organized by The Institute of Mathematical Sciences, Chennai.
- 8.International Conference of Dynamical System: Theory and Application (11th -14th January 2012) organized by department of mathematics, Jadavpur University, and Biomathematical society of India.
- 9.Probability and Discrete Mathematics in Mathematical Biology (14th March–10th May 2011). Institute for Mathematical Sciences, National University of Singapore. Workshop Attended: Discrete Mathematics and Probability in Population Biology and Genetics 21st March-1st April, 2011 Discrete Mathematics and Probability in Networks and Population Biology 2nd May-13th May. Delivered talk on "Chaos in Ecological Systems".
- 10.National Workshop on Nonlinear Dynamical Systems (4th-8th July 2011). Department of Mathematics, National Institute of Technology, Durgapur.
- 11. Workshop & Symposium on Mathematical Ecology (7_{th} - 14_{th} December 2010) organized by Indian Institute of Science Mathematics Initiative, IISER Kolkata & DST Centre for Mathematical Biology
- 12. National Conference on Theoretical Biology and Biomathematics" (15th –16th December 2010) organized by
- "Biomathematical Society of India" & "Centre for Mathematical Biology & Ecology, Department of Mathematics, Jadavpur University"

Scientific Activities & Achievements

- 1. Nominated as a member of Jury Panel to evaluate the research projects of Avishkar Research Convention of university of Mumbai for the final round selection (2017).
- 2.Indian Statistical Institute Teaching Assistant (2010-2011) Course: Analysis of Discrete Data, M.Stat II(Designing Assignment and half-semester teaching)
- 3. Qualified as Junior Research Fellow in "Council of Science & Industrial Research (CSIR) -NET", June-2009.
- 4.Best paper presentation in Mathematics session in the International conference **India Biodiversity Meet, 2013.** Indian Statistical Institute, Kolkata, India. **Title:** Behavior of Nonlinear Birth-Death Process under Quasi Equilibrium with Application in Growth Study.
- 5.Member of "Biomathematical Society of India" (http://biomathsociety.in/)
- 6.Member of the organizing committee of the Annual International event "India Biodiversity Meet" organized by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata.

7.One of the core committee members of the "Workshop on Statistical Methods and R Programming" organized annually by the Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata.

Travel grants

1.International Travel Grant International Travel Grant from National Board of Higher Mathematics to attend AARMS Summer school on Statistical Learning in Dalhousie University, Halifax, Canda APRIL 2014
2.International Travel Grant Travel grant from Institute of Mathematical Sciences, National University of Singapore to attend international workshop and two months research collaboration June 2012

Education

1.University of Calcutta

PhD in Applied Mathematics (2015), Research work carried out as Research fellow in Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata.

2.Indian Institute of Technology Bombay (2006-2008)

M.Sc. in Applied Statistics & Informatics (First Division)

3.Presidency College, Kolkata (2003-2006)

B.Sc. in Mathematics (Honours) (First Division)

4. Chakdaha Ramlal Academy, Chakdaha (2001-2003)

West Bengal Council of Higher Secondary Education (First Division)

5. Rautari High School, Rautari, Nadia (2001)

West Bengal Board of Secondary Education (First Division)

PhD Courses completed

Ecological Statistics, Distribution theory and Statistical Inference, Advanced Stochastic Process, Metric Topology and Complex Analysis, Statistical Methodology, Computer Programming, Population Dynamics, Mathematical Modelling of Biological Systems.

Other Skills

Knowledge of Research Methodologies Statistical Software: Good knowledge in R Writing and Presenting research reports

Mathematical Software: Matlab, Latex, Mathematica

Other Information

Permanent Address

Village – Bramhapara, P.O. – Simurali,

PS-Chakdaha District - Nadia,

PIN – 741248, West Bengal, India

Office Address

Department of Mathematics Institute of Chemical Technology, Mumbai

Nathalal Parekh Marg, Mumbai-400019

Language

Fluent Speaking and Writing in Bengali and English. Moderate speaking and writing in Hindi.

Summary of PhD research

Growth curve models serve as the mathematical framework for the quantitative studies of growth in many areas of applied science. In Ecology, the density dependent models refer to the population dynamics, where, growth rate depends on some functions of population size. In this thesis we have extended the density dependent model based on the ecological foundation of cooperation. To validate this idea, we used Global Population Dynamics Database, a vast repository of population time series data containing more than 5000 time series, and Sibly's (Science, 2005) framework of studying growth regulation of animal populations. The proposed methodology helps better explain the extinction risk of species and equip us with a good conservation management tool. In

natural populations, the stochastic set up is more appropriate to take into account randomness. We use a more general form of birth and death process to extend the cooperation model under stochastic environment. We propose a new method for approximating the moments of equilibrium distribution that is valid for both integer and non-integer values of model parameters. Apart from cooperation, memory is also an essential behavioural issue and is an integrated part of living organisms. We extend the density dependent and independent models to include the effect of memory by using fractional calculus. In the next phase of thesis, we introduce the concept of time covariate model, where growth rate is a function of time and discuss its utility to explain a real life biological experiment and study related statistical inference problems. The pillar of the growth models is the relative growth rate (RGR). This RGR is empirically estimated by Fisher (1921)'s average RGR. This average RGR is growth law invariant metric based on the assumption of exponential growth between two consecutive time points. We propose the concept of interval specific rate parameter that captures the proximity of a given data with respect to underlying model even for short time interval. The proposed measure can be used as a more appropriate model selection criterion than the existing summary measures of goodness of fit. Potential future directions are given at the end of the thesis.

References

1.Dr. Sabyasachi Bhattacharya, PhD, Associate Professor (sabyasachi@isical.ac.in) ((+91)-9433897120)

Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata

2.Dr. Santanu Ray, Professor, Visva Bharati University Ecological Modelling Laboratory, Department of $Zoology(\underline{sray@visva-bharati.ac.in})$ (+91 – 9433157701)

Homepage: https://sites.google.com/site/rakeshkunduvb/people-1/prof-santanu-ray

3.Dr. Joydev Chattopadhyay, Professor, PhD, FASc (joydev@isical.ac.in) ((+91)-9830546490)

Agricultural and Ecological Research Unit, Indian Statistical Institute, Kolkata Homepage: http://www.isical.ac.in/~joydev/main.html

4.Dr. Ayanendranath Basu, Professor (<u>ayanbasu@isical.ac.in</u>) (+91 33 2575 2806) Interdisciplinary Statistical Research Unit (ISRU), Indian Statistical Institute, Kolkata

Homepage: http://www.isical.ac.in/~ayanbasu/

Declaration

I hereby declare that the above mentioned information is correct to the best of my knowledge and belief.

Place: Mumbai

Dr. Amiya RanjanBhowmick

9. Feedback of the Participants

Sr.	Question	Response				
No.		Satisfactory	Average	Good	Very Good	Excellent
1	Enjoyed the workshop on the whole	0	1	3	25	18
2	The content was as described in publicity materials	0	1	6	19	21
3	I will recommend this workshop to others	0	1	3	17	26
4	The program was well paced within the allotted time	1	3	6	19	18
5	Instructor has good communication skills	0	1	1	9	36
6	The material was presented in an organized manner	2	0	2	14	29
7	The instructor is highly knowledgeable on the topic	1	0	1	4	41
8	I would be interested in attending advanced workshop on this subject	0	1	4	10	32

10. Receipts and Expenditure Statement

Sr. No.	Income Head	Amount
1	Total number of Participants=52*500	Rs.26,000/-
2	BMESI Sponsorship	Rs.7,500/-
	Total Amount	Rs. 33,500/-
Sr. No.	Expenditure Head	Total Cost
1	Food- Tea+Lunch+Tea	Rs.15,500/-
2	Speaker Remuneration – Rs.5000/day	Rs.10,000/-
3	Transport Charges	Rs.1715/-
4	Printing (Brochure, Poster, Flex, Certificate)	Rs.3700/-
5	Stationary- (Notepad, Pen, Folder)	Rs.650/-
6	Stage Decoration + Momentum	Rs.2677/-
	Total Expenditure	Rs. 34,242/-

Sr. No	Name	SBI Receipt Number
1.	Abhisekh Mankar	DU87332384
2.	Aeshaben Gandhi	DU87460499
3.	Afifa Juhi Khan	DU7341058
4.	Akhidev Devdas	DU87342252
5.	Annesha Moza	DU87521888
6.	Anisha Jagdale	DU87359933
7.	Aparna Lakhe	DU87431558
8.	Arbacheena Bora	DU87748960
9.	Ashraf Firoz Shaikh	DU87472316
10.	Ashwini Tidke	DU87749945
11.	Chetan Ippar	DU87350923
12.	Claire Vania	DU87332866
13.	Deepa Kale	DU87451602
14.	Ganesh Teke	DU87360793
15.	Gomati Nandekar	DU87461861
16.	Harshal Raghotre	DU87521277
17.	Jjismi Easow	DU87345303
18.	Jyoti Jethe	DU87405318
19.	Kanan Dombgare	DU87458808
20.	Khushboo Dhuri	DU87484570
21.	Krutika Jadhav	DU87344938
22.	Krutika Jagdale	DU87428367
23.	Manasi Sawant	DU87431013
24.	Mansi Gaikwad	DU87457480
25.	Mayur Ghangale	DU87565322
26.	Mayuri Mulay	DU87334048
27.	Mohini Mehta	DU87460888
28.	Nikita Panchal	DU87491838
29.	Nikita Panyam	DU87430937
30.	Nishant Patil	DU87490143
31.	Nischita Moundekar	DU87475907
32.	Parth Gamare	DU87463934
33.	Prajakta Mhatre	DU87456429
34.	Priya Gaikar	DU87349351
35.	Pumpy Kathar	DU87344507
36.	Rajpal Jagtap	DU87344510
37.	Salik Khan	DU87483512
38.	Shashikant Rathod	DU87748204
39.	Sharda Tondare	DU87722718
40.	Shrikant Patil	DU87520999
41.	Shruthy Seshadrinathan	DU87421284
42.	Shruti Bajpai	DU87748793
43.	Shruti Gadkari	DU87344125
44.	Shweta Chaudhari	DU87749649
45.	Shweta Singh	DU87427758
46.	Siddhant Ojha	DU87520496
47	Simanjay Singh	DU87565319

48.	Sneha Bansod	DU87482961
49.	Sushma Bhat	DU87489752
50.	Sushma Nair	DU87340091
51.	Tejasvini Shendge	DU97491842
52.	Uday Shirsekar	DU87341310