Soudeep Deb

CURRENT POSITION
Assistant Professor
Decision Sciences Area
Indian Institute of Management, Bangalore, India.

Webpage: soudeepd.github.io
Phone: +91 80-26993387
E-mail: soudeep@iimb.ac.in

CITIZENSHIP India

RESEARCH INTERESTS Time series data, Spatial statistics, Spatio-temporal modeling, Clustering and classification, Sports analytics, Application of time series and spatial statistics in finance and other disciplines.

EXPERIENCE

Indian Institute of Management Bangalore, KA, India. Mar 2020 - Present

- Assistant Professor, Young Faculty Research Chair, Decision Sciences Area.

NBC Universal Media, LLC., New York, NY, USA. Sep 2018 - Feb 2020

- Senior Lead Data Scientist, Decision Sciences Division.

EDUCATION

University of Chicago, Chicago, IL, USA.

Ph.D., Statistics Aug 2018

- Thesis: Irregular spaced random field, Spatio-temporal data and Clustering of time series
- Advisor: Dr. Wei Biao Wu
- Other committee members: Dr. Ruey S. Tsay and Dr. Michael L. Stein

Indian Statistical Institute, Kolkata, WB, India.

Master of Statistics (M. Stat.) May 2013

- First Division with Distinction
- Specialization: Mathematical Statistics and Probability
- Dissertation: Association analysis for identifying rare genetic variants
- Advisor: Dr. Saurabh Ghosh

Bachelor of Statistics (B. Stat.) May 2011

- First Division with Distinction

TEACHING

Course instructor, at Indian Institute of Management, Bangalore:

- Multivariate Statistics (FPM/PhD) Term 3 of 2021
- Decision Sciences I (PGP/MBA) Term 1 of 2021
- Statistical Inference (FPM/PhD) Term 2 of 2020
- Decision Sciences I (PGP/MBA) Term 1 of 2020

Course instructor, at University of Chicago:

- Introductory Statistics at Chicago Academic Achievement Program Summers of 2015, 2017
- Statistical Models and Methods I Winter 2015

RESEARCH GRANTS

Research Seed Grant from IIM Bangalore, India:

- Project: New techniques to analyze categorical and discrete time series data
- Amount: INR 250,000 for two years (August 2021 to July 2023)

Young Faculty Research Grant from IIM Bangalore, India:

- Amount: INR 900,000 for three years (March 2020 to February 2023)

Publications

1. **Deb, S.**, Deb, S. (2022). An ensemble method for early prediction of dengue outbreak. Journal of the Royal Statistical Society Series A, 185(1), 84-101.
2. Rawat, S., **Deb, S.** (2021). A spatio-temporal statistical model to analyze COVID-19 spread in the USA. Accepted in the Journal of Applied Statistics.
3. **Deb, S.** (2021). Analyzing airlines stock price volatility during COVID-19 pandemic through internet search data. Accepted in the International Journal of Finance & Economics.
4. **Deb, S.**, Tsay, R. S. (2019). Spatio-temporal Models with Space-time Interaction and Their Applications to Air Pollution Data. Statistica Sinica, 29, 1181-1207.
5. **Deb, S.**, Dey, D. (2019). Spatial Modeling of Shot Conversion in Soccer to Single out Goalscoring Ability. Journal of Sports Analytics, 5(4), 281-297.
6. **Deb, S.** (2019). VAR Model Based Clustering Method for Multivariate Time Series Data. Journal of Mathematical Sciences, 237(6), 754-765.
7. Prickett, K.C., Guiterrez, C., **Deb, S.** (2019). Family Firearm Ownership and Firearm-related Mortality among Young Children: 1976-2016. Pediatrics, 143(2), e20181171.
8. Chazin, H., **Deb, S.**, Falk, J., Srinivasan, A. (2019). New Statistical Approaches to Intra-individual Isotopic Analysis and Modelling of Birth Seasonality in Studies of Herd Animals. Archaeometry, 61(2), 478-493.
9. **Deb, S.**, Pourahmadi, M., Wu, W. B. (2017). An Asymptotic Theory for Spectral Analysis of Random Fields. Electronic Journal of Statistics, Vol. 11, No. 2, p. 4297-4322.
10. Zechner, C., **Deb, S**., Koeppl, H. (2013). Marginal Dynamics of Stochastic Biochemical Networks in Random Environments. In Control Conference (ECC), 2013 European, p. 4269-4274, IEEE.
11. Ghosh, S., **Deb, S.** (2013). A Clustering Approach for Mapping Rare Variants Based in Mutual Association. Human Heredity, Vol. 76, No. 2, pp. 98-98.

Submitted articles

1. **Deb, S.**, Karmakar, S. (2022). A novel spatio-temporal clustering algorithm with applications on COVID-19 data from the United States. Under review. Pre-print available on request.
2. **Deb, S.**, Jana, K. (2022). Nonparametric quantile regression for time series with replicated observations and its application to climate data. Under review. Pre-print: https://arxiv.org/abs/2107.02091.
3. Bag, S., Gupta, K., **Deb, S.** (2022). A review and recommendations on variable selection methods in regression models for binary data. Under review. Pre-print: https://arxiv.org/abs/2201.06063.
4. **Deb, S.** (2022). Analyzing count data using a time series model with an exponentially decaying covariance structure. Under review. Pre-print: https://arxiv.org/abs/2004.03130.
5. Gupta, K., Krishnamurthy, V., **Deb, S.** (2022). What elements of the opening set influence the outcome of a tennis match? An in-depth analysis of Wimbledon data. Pre-print available on request.
6. Paul, M., Roy, R., **Deb, S.** (2022). Effect of influence in voter models and its application in detecting frauds in an election. Under review. Pre-print available on request.
7. Mareeswaran, M., Sen, S., **Deb, S.** (2021). New methods of structural break detection and an ensemble approach to analyze exchange rate volatility of Indian rupee during COVID-19. Under review. Pre-print available on request.
8. **Deb, S.** (2021). A mathematical take on the competitive balance of a football league. Under revision. Pre-print: https://arxiv.org/abs/2102.09288.
9. Nahata, S., **Deb, S.** (2021). A machine learning approach to analyze the effect of situational and player-dependent features on converting freekicks in soccer. Under review.
10. **Deb, S.**, Majumdar, M. (2020). A time series method to analyze incidence pattern and estimate reproduction number of COVID-19. Under review. Pre-print: https://arxiv.org/abs/2003.10655.
11. Majumdar, M., Banerjee, M., Sengupta, J., **Deb, S.**, Jana, C. K., Roy, B. K. (2020). Prevalence and spectrum of diabetic peripheral neuropathy and its correlation with insulin resistance – An experience from eastern India. Pre-print: https://www.medrxiv.org/content/10.1101/2020.04.12.20056150v1.

Seminars

1. A new classification method for multivariate time series data. 14th International Conference of the ERCIM WG on Computational and Methodological Statistics, London, UK. Dec 2021
2. Impact of global warming on rainfall: A Spatio-temporal study. RSS International conference 2021 (virtual), Manchester, UK. Sep 2021
3. Analyzing count data using a time series model with an exponentially decaying covariance structure. 2021 Australian and New Zealand Statistical Conference (virtual), Australia. Jul 2021
4. Spectral density based clustering method for spatio-temporal datasets. 4th International Conference on Econometrics and Statistics, EcoSta 2021 (virtual), Hong Kong. Jun 2021
5. A Mathematical Take on the Competitive Balance of a Football League. Guest lecture (virtual), Mathematics of Sports, Stanford University, Stanford, USA. May 2021
6. New Methods of Clustering Time-series Data and its Applications. Colloquium, Indian Statistical Institute, Bangalore, India. Nov 2019
7. Spatio-temporal Models with Space-time Interaction and Their Applications to Air Pollution Data. Statistics colloquium, Northern Illinois University, Dekalb, USA. Dec 2017
8. VAR Model Based Clustering Method for Multivariate Time Series Data. XXXIV. International Seminar on Stability Problems for Stochastic Models, Debrecen, Hungary. Aug 2017
9. Spatio-temporal Models with Space-time Interaction and Their Applications to Air Pollution Data. NBER/NSF Time Series Conference, New York, USA. Sep 2016
10. Hydropower Construction and Deforestation in the Tapajós River Basin: The Effect on Water Balance. Symposium on deforestation, Ministry of Environment, Brasília, Brazil. Aug 2016
11. Estimating Genetic Relationship between Random Individuals from Genome Sequence Data. Young Statisticians Conference, Melbourne, Australia. Feb 2013
12. Estimating Genetic Relationship between Random Individuals from Genome Sequence Data. Conference on Contemporary Issues and Applications in Statistics, Kolkata, India. Jan 2012
13. Estimating Genetic Relationship between Random Individuals from Genome Sequence Data. D. Basu Memorial Award Lecture Series, Indian Statistical Institute, Kolkata, India. Sep 2011

Peer review Services

Reviewer for the following journals:

- Biometrics
- Electronic Journal of Statistics
- IIMB Management Review
- Indian Journal of Science and Technology
- Journal of Advances in Management Research
- Journal of Behavioral and Experimental Finance
- Journal of Sports Analytics
- Journal of the Royal Statistical Society: Series C
- Linear Algebra and its Applications
- Sankhya A and Sankhya B
- Statistics and Probability Letters

Honors

University of Chicago:

- International House Ralph W. Nicholas Fellowship Award — 2017-18
- Graduate Council Travel Fund Award — 2017
- Senior Consultant, Department of Statistics — 2016-17
- Runner-up for Department of Statistics Consulting Award — 2016
- Nominated for Best Teaching Assistant in Physical Sciences Division — Winter 2014

Other Awards:

- Kishore Vaigyanik Protsahan Yojana scholarship, Indian Institute of Science — 2007 to 2013
- Selected for International Mathematical Olympiad Training Camp, India — 2007 & 2008

Other Experience

The Alan Turing Institute, London, United Kingdom — Dec 2017

- Position: Delegate for the Data Study Group.
- Project: Geospatial time-series analyses to predict demand for a global satellite communications network.

Instituto de Pesquisa Ambiental do Amazônia, Brasília, Brazil — Jun - Aug 2016

- Position: Summer fellow.
- Project: Hydropower Construction and Deforestation in the Tapajós River Basin: Linking Forest Cover to Changes in Water Balance.

Eidgenossische Technische Hochschule (ETH), Zurich, Switzerland — May - Jul 2013

- Position: Summer research intern.
- Project: Moment-Closure Approximations for Mass-action Models in Chemical Kinetics.

Eidgenossische Technische Hochschule (ETH), Zurich, Switzerland — Jun - Jul 2012

- Position: Summer research intern.
- Project: Marginal Dynamics of Stochastic Biochemical Networks in Random Environments.

Ministry of Statistics and Programme Implementation, Govt. of India — May 2012

- Position: Team member
- Project: Forecasting of Foreign-tourist Arrivals in India.

Johns Hopkins University, Baltimore, United States of America — May - Jul 2011

- Position: Summer research intern
- Project: Estimating Relationship between Random Individuals from Genome Sequence Data.

Skills

Technical strength:

- Proficient: R, MATLAB, LaTeX, Microsoft Office
- Working knowledge: Python, SQL, C.

Languages:

- Fluent in reading, writing, speaking: English, Bengali, Hindi.
- Basic reading and speaking: Portuguese.