Computational Statistics Research Laboratory

We practice data science, we train data scientists.

Description:

The main interest in the laboratory is the development of statistical theory and methods with the aid of computing. The group specifically works on two different but complementing broad areas of empirical nonparametrics and high dimensional data visualization and analysis. Nonparametric and robust methods are used to perform structural inference in spatio-temporal and additive models and in model-based estimation in survey sampling. We also consider some stylized facts related to these models like structural change, spatial heterogeneity, volatility, clustering, complex dependence structure (space, time, space-time interactions) among others. On high dimensional data, we study sparsity in relation to visualization and model selection, clustering, robustness, and other problems and issues usually encountered in data mining and activities related to business intelligence.

The applications of the theory and methods developed within the laboratory have found its applications in areas that inspired the formulation of these problems including: genomics; biostatistics; data mining in large databases (including ***big data***); analytics; consumer behavior modeling; estimation in complex survey designs; monitoring and impact assessment; risk analytics; among others.

Current Members:

Erniel B. Barrios
Martin Augustine B. Borlongan
Wendell Q. Campano
John Carlo P. Daquis
John Eustaquio
Ivy Iris M. Gauran
Joseph Ryan G. Lansangan
Michael Daniel C. Lucagbo
Angela D. Nalica
Paolo Victor T. Redondo
Genelyn Ma. F. Sarte
Michael Van B. Supranes
Kevin Carl P. Santos
Stephen Jun V. Villejo

Graduate Students:

PhD (Statistics)

Ivy Corazon L. Ancog (high dimensional switching meta-regression models)
Wendell Q. Campano (spatio-temporal modeling of time series count with volatility)
Carlito Daarol (longitudinal data modeling with copula)
John Carlo P. Daquis (longitudinal count data)
Calixto Elnas (spatiotemporal switching models)
Aristotle Magallanes (spatio-temporal modeling of epidemics)
Kevin Carl P. Santos (cognitive diagnostic modeling)

MS (Statistics)

Johann Sebastian Claveria (testing for overdispersion in poisson regression)
Ronrick Daano (multilevel model with structural change)
Mary Ann S. Estoy (quantile regression of clustered data)
Virgilio Llenansares (nonparametric transfer function models)
Michelle Macapinlac (high dimensional autoregression)
Daniel R. Raguindin (semiparametric probit model for high dimensional clustered data)
Nenita Rebarter (sampling from skewed population)
Karl Anton Retumban (robust estimation of a nonlinear model)
Paolo Redondo (estimation of poisson autoregression model)
Phoebe Sombillon (dynamic spatio-temporal models)
Michael Van B. Supranes (bi-clustered high dimensional data)

Master of Statistics

Joseph L. Arbizo (forecasting count time series)
Damaris Yarcia (contagion and business cycles)

Awards:

Erniel B. Barrios – Gawad Chancellor sa Natatanging Guro 2010; UP Scientist III
Ivy Iris M. Gauran – Jan Tinbergen Award 2013, International Statistical Institute
Joseph Ryan G. Lansangan – UP Scientist I
Kevin Carl P. Santos – Cochran-Hansen Prize 2015, International Statistical Institute

Publications:

1. C. Manalaysay and **E. Barrios** (2016), Semiparametric Principal Components Poisson Regression on Clustered Data, forthcoming in *Communications in Statistics-Simulation and Computing*.
2. R. Guilatco and **E. Barrios** (2016), Nonparametric Estimation of a Switching Regression Model, forthcoming in *Communications in Statistics-Simulation and Computing*.
3. **I. Gauran** and **E. Barrios** (2016), Nonparametric Modeling of Clustered Customer Survival Data, forthcoming in *Communications in Statistics-Simulation and Computing*.
4. **K. Santos** and **E. Barrios** (2016), Improving Predictive Accuracy of Logistic Regression Model Using Ranked Set Samples, forthcoming in *Communications in Statistics-Simulation and Computing*.
5. M. Poblador and **E. Barrios** (2016), Control Chart for Monitoring Autocorrelated Process with Multiple Exogenous Inputs, forthcoming in *Communications in Statistics-Simulation and Computing*.
6. A. Asaad and **E. Barrios** (2015), Nonparametric Bootstrap Test in a Multivariate Spatial-Temporal Model: A Simulation Study, *The Philippine Statistician, 64(1):1-16.*
7. **E. Barrios** (2015), Robustness, Data Analysis, and Statistical Modeling: The First 50 Years and Beyond, *Communications for Statistical Applications and Methods, 22(6):543-556.*
8. **S. Villejo** (2015), Modelling Rice Yield in the Philippines Using Dynamic Spatio-temporal models, *The Philippine Statistician, 64(2):43-58*.

9. **S. Villejo** (2015), Classification and Prediction of Suicidal Tendencies of the Youth in the Philippines: An Empirical Study, *The Philippine Statistician, 64(1):31-52*.
10. **J Eustaquio**, D Mapa, M Mindanao, N Paz (2015), Forecasting Time-varying Correlation using the DCC Model, *The Philippine Statistician, 64(1):53-74*.
11. **A. Nalica, I. Gauran** (2014), Modeling Clustered Survival Data with Cured Fraction, *The Philippine Statistician, 63(2):1-8*.
12. **K. Santos** (2014), Modeling Zero-Inflated Clustered Count Data: A Semiparametric Approach, *The Philippine Statistician, 63(1): 21-32*.
13. **M. Supranes** (2014), Design Strategies in Fitting a Nonlinear Model, *The Philippine Statistician, 63(1): 43-58*.
14. **A. Nalica, J. Lansangan** (2014), Identifying Influencers of Consumer Activity: A Case Study in Predictive Analytics, *The Philippine Statistician, 63(1): 93-102*.
15. R. Veron Cruz and **E. Barrios** (2014), Estimation Procedure for a Multiple Time Series Model, *Communications in Statistics-Simulation and Computing*, 43(10):2415-2431.
16. J. Umali and **E. Barrios** (2014), Nonparametric Principal Components Regression, *Communications in Statistics-Simulation and Computing*, 43 (7):1797-1810.
17. **I. Gauran,** S. Poblador, S., (2013), Classification of Congenital Hypothyroidism using Articial Neural Networks, *The Philippine Statistician*, 62(2): 1-11.
18. **Campano, W.**, Tadlas, R., (2013), Visual Exploration of Climate Variability, *The Philippine Statistician*, 62(2): 101-111.
19. **J. Daquis,** (2013), Nonparametric Transfer Function Models with Localized Temporal Effect, *The Philippine Statistician*, 62(1): 1-14.
20. **J. Lansangan** (2013), Sparse Principal Component Regression, *The Philippine Statistician*, 62(1): 33-50.
21. J. Guarte and **E. Barrios** (2013), Nonparametric Hypothesis Testing in a Spatial-Temporal Model: A Simulation Study, *Communications in Statistics-Simulation and Computing*, 42(1):153-170.
22. M. Mancenido and **E. Barrios** (2012), An AR-Sieve Bootstrap Control Chart for Autocorrelated Process Data, *Quality and Reliability Engineering International*, 28:387-395.
23. **J. Daquis,** M. Laus, N. Supnet (2012), Assessing Strength of Seasonality Through Sample Entropy: A Simulation Study, *The Philippine Statistician*, 61(1): 21-34.
24. S. Poblador, **I. Gauran** (2012), Sampling with Probability Proportional to Aggregate Size using Nonparametric Bootstrap in Estimating Total Production Area of Top Cereals and Root Crops Across Philippine Regions, *The Philippine Statistician*, 61(1): 87-108
25. **K. Santos,** C. Castillo, R. de Jesus, N. Telan, C. Vidal (2012), Nonparametric Bootstrap Estimation of the Population Ration Using Ranked Set Sampling, *The Philippine Statistician*, 61(2): 53-66.
26. **W. Campano** (2012), Robust Methods in Time Series Models with Volatility, *The Philippine Statistician*, 61(2): 83-102.
27. E. Santos and **E. Barrios** (2012), Nonparametric Decomposition of Time Series Data with Inputs, *Communications in Statistics-Simulation and Computing*, 41:1693-1710.
28. **E. Barrios** and **J. Lansangan** (2012), Forecasting Customer Lifetime Value: A Statistical Approach, *Philippine Management Review*, 19: 23-34.

29. **K. Santos** and J. Salagubang (2011), Investigating the Efficiency of Stratified Ranked Set Sampling Using Nonparametric Bootstrap Estimation, *The Philippine Statistician*, 60:15-30.
30. **J. Lansangan** (2011), A Dose of Business Intelligence: Data Mining, *The Philippine Statistician*, 60:125-128.
31. L. Santos and **E. Barrios** (2011), Small Sample Estimation in Dynamic Panel Data Models: A Simulation Study, *Open Journal of Statistics*, 1(2):58-73.
32. **W. Campano** and **E. Barrios** (2011), Robust Estimation of a Time Series Model with Structural Change, *Journal of Statistical Computing and Simulation*, 81(7):909-927.
33. R. Bastero and **E. Barrios** (2011), Robust Estimation of a Spatiotemporal Model with Structural Change, *Communications in Statistics- Simulation and Computation*, 40(3):448-468.
34. **E. Barrios** (2011), Bootstrap Methods, *The Philippine Statistician*, 60:129-132.
35. C. Mina and **E. Barrios** (2010), Profiling Poverty with Multivariate Adaptive Regression Splines, *Philippine Journal of Development*, 69(2):55-97.
36. M. Chernick, W. Gonzalez-Manteiga, R. Crujeiras, and **E. Barrios** (2010), Bootstrap Methods, in M. Lovric, ed., International Encyclopedia of Statistical Sciences, N.J: Springer. pp. 169-174.
37. C. Dumanjug, **E. Barrios**, and **J. Lansangan** (2010), Bootstrap Procedures in a Spatial-Temporal Model, *Journal of Statistical Computing and Simulation*, 80(7):809-822 .
38. **E. Barrios** and L. Lavado (2010), Spatial Stochastic Frontier Models, Discussion Paper Series No. 2010-08, *Philippine Institute for Development Studies*, May 2010.
39. **J. Lansangan** and **E. Barrios** (2009), Principal Components Analysis of Nonstationary Time Series Data, *Statistics and Computing*, 19(2): 173-187.
40. R. Lavado and **E. Barrios** (2008), Spatial-Temporal Dimensions of Efficiency Among Electric Cooperatives in the Philippines, Discussion Paper Series No. 2008-29, *Philippine Institute for Development Studies*, November 2008.
41. **E. Barrios** and **G. Sarte** (2008), Monitoring Sustainable Agriculture in Southeast Asia, *International Journal of Sustainable Development and World Ecology*, 15:95-102..
42. **E. Barrios** (2007), Model-Based Predictive Estimation with Coverage Error, *Bulletin of the International Statistical Institute*, LXII:4398-4401.
43. **A. Nalica** and **E. Barrios** (2007), Approaches in Forecasting Cereal Production, *The Philippine Statistician*, 56:85-102.
44. **E. Barrios** (2007) Spatial Effect in the Efficient Access of Rural Development, ADBI Discussion Paper No. 65,*Asian Development Bank Institute*, Tokyo, Japan, May 2007(http://www.adbi.org/files/dp65.rural.development.efficieny.access.pdf).
45. **O. Landagan** and **E. Barrios** (2007), An Estimation Procedure for a Spatial-Temporal Model, *Statistics and Probability Letters*, 77:401-406.
46. **E. Barrios** and K. Komoto (2006), Some Approaches in the Construction of a Sustainable Development Index for The Philippines, *International Journal of Sustainable Development and World Ecology*, 13:277-288.
47. J. Guarte and **E. Barrios** (2006), Estimation Under Purposive Sampling, *Communications in Statistics- Simulation and Computation*, 35:277-284.
48. B. Fermin and **E. Barrios** (2002), Autologistic Models with Autocorrelated Errors for an Aquatic Phenomenon, *The Philippine Statistician*, 51:43-52.

Papers in preparation/under review:

1. **J. Lansangan** and **E. Barrios**, Simultaneous Dimension Reduction and Variable Selection in Modeling High Dimensional Data, *for revision.*
2. **J. Eustaquio** and **E. Barrios**, Nonparametric Hypothesis Testing in Clustered Survival Model, *for revision.*
3. **M. Lucagbo** and **E. Barrios**, Estimation of Isotonic Spatio-Temporal Model with Clustering, *under review.*
4. M. Ang and **E. Barrios**, Cross-Country Comparison of Maternal Mortality Ratio, *under review.*
5. **E. Barrios** and **J. Lansangan**, Mining Insights on the Quality of Life Across Countries: Lessons from Official Statistics, *in preparation.*
6. M. Esmenda and **E. Barrios**, Robust Estimation of Multilevel Model with Structural Change, *under review.*
7. **S. Villejo**, **E. Barrios**, **J. Lansangan**, Robust Estimation of a Dynamic Spatio-Temporal Model with Structural Change, *under review.*
8. **E. Barrios,** R. Olivares, S. Data, Robust Estimation in Clustered Multiple Time Series with Structural Change, *under review.*
9. D. Del Prado and **E. Barrios**, Small Area Estimation Using Spatio-Temporal Mixed Model, *in preparation.*
10. M. Torres and **E. Barrios**, Sparse Nonparametric Discrete Choice Model for High Dimensional Data, *in preparation.*
11. **M. Borlongan**, **E. Barrios**, **J. Lansangan**, Robust Simultaneous Confidence Interval Estimation of Principal Component Loadings, *in preparation.*
12. V. Alao, **E. Barrios**, **J. Lansangan**, Semiparametric Estimation of Mixed Analysis of Covariance Model, *in preparation.*

Dissertation Completed:

1. **Del Prado, PhD, 2014**: Spatio-Temporal Mixed Model in Small Area Estimation of Data from Rotated Panel Surveys.
2. **Lansangan, PhD, 2014**: Simultaneous Dimension Reduction and Variable Selection in Modeling High Dimensional Data.
3. **Asaad, PhD, 2013**: Nonparametric Bootstrap Inference in a Multivariate Spatial-Temporal Model.
4. **Guarte, PhD, 2009:** Nonparametric Bootstrap Inference in a Spatial-Temporal Model.

Thesis Completed:

1. **Borlongan, MS, 2016**: Robust Simultaneous Confidence Interval Estimation of Principal Component Loadings.
2. **Alao, MS, 2015**: Semiparametric Estimation of Mixed Analysis of Covariance Model.
3. **Saavedra, MS, 2015**: Robust Inference in Regression Models.
4. **Villejo, MS, 2015**: Robust Estimation of a Dynamic Spatio-Temporal Model with Structural Change.
5. **Ramos, MS, 2014**: Estimation of Multiple Time Series with Volatility.
6. **Abeto, MOS, 2014**: SPCR-Based Control Chart for Autocorrelated Processes with High Dimensional Inputs.

7. **Oberos, MOS, 2014**: Estimation in Repeatitive Surveys: The Case of Employment Data.
8. **Eustaquio, MS, 2014**: Nonparametric Hypothesis Testing in Clustered Survival Model.
9. **Tobias, MS, 2014**: A Comparison of Parametric and Nonparametric Estimation of Multiple Input Transfer Function Model.
10. **Olivares, MS, 2014**: Robust Estimation of Clustered Multiple Time Series with Structural Change
11. **Data, MS, 2014**: Estimation of Independent Multiple Time Series Model with Structural Change.
12. **Paragas, MS, 2014**: Estimation of Multilevel Model with High Dimensional Data.
13. **Paglinawan, MOS, 2013**: Comparison of Regression Estimator and Ratio Estimator: A Simulation Study.
14. **Torres, MS, 2013**: Sparse Nonparametric Discrete Choice Model for High Dimensional Data.
15. **Esmenda, MS, 2013**: Robust Estimation of Spatio-Temporal Multilevel Model with Structural Change.
16. **Lucagbo, MS, 2013**: Estimation of Isotonic Spatio-Temporal Model with Clustering.
17. **Dimapilis, MOS, 2012**: AR-Sieve-Based Prediction Interval For Sustainable Development Index.
18. **Ofina, MS, 2012**: Autologistic Spatial-Temporal Modeling.
19. **Poblador, MS, 2012**: A Control Chart Design for Monitoring Autocorrelated Process with Multiple Exogenous Inputs Under Model Uncertainty.
20. **Gauran, MS, 2012**: Nonparametric Modelling of Clustered Survival Data.
21. **Lavina, MOS, 2012**: Assessing the Impact of Short-Term Shocks to Inflation Forecasts
22. **Veron Cruz, MS, 2012**: Estimation of a Model for Multiple Time Series.
23. **Mancenido, MOS, 2011:** An AR-Sieve Bootstrap Control Chart for Autocorrelated Process Data
24. **Manalaysay, MS, 2011:** Semiparametric Principal Component Poisson Regression on Clustered Data
25. **Estiaga, MOS, 2011:** Bootstrapping Penalty Analysis in Sensory Evaluation of Pizza Products
26. **Santos K, MS, 2011:** Predictive Accuracy of Fitted Logistic Regression Model Using Ranked Set Samples
27. **Salagubang, MS, 2011:** Outlier Detection in High Dimensional Data In The Context of Clustering
28. **Esquierres, MS, 2011:** Semiparametric Spatial-Temporal Model
29. **Mantuano, MS, 2011**: Comparison of Predictive Ability of Two Competing ARMA Models Via Cross-Spectral Analysis
30. **Gonzales, MOS, 2011**: Credit Card Application Scoring System
31. **Abrigo, MOS, 2011**: Estimating Counterfactual Schooling Outcomes Using Inverse Propensity Score Reweighting.
32. **Umali, MS, 2010:** Nonparametric Principal Components Regression
33. **Devera, MS, 2010:** Semiparametric Poisson Regression for Clustered Data (Finalist DOST-PCASTRD Outstanding Thesis in Advanced Science and Technology, 2010)
34. **Santos E, MS, 2010:** Decomposition of Multicollinear and Time Series Data Using Backfitting and Additive Models

35. **Daquis, MS, 2010:** Nonparametric Transfer Function Models with Localized Temporal Effects.
36. **Santos J, MS, 2010:** Robust Inference in Semiparametric Spatiotemporal Models
37. **Matammu, MS, 2010:** Testing for Presence of Structural Change and Spatial Heterogeneity in Spatiotemporal Models with Temporal Volatility
38. **Giron, MS, 2010:** Discrete Choice Generalized Additive Models Using High Dimensional Data
39. **Kwong, MS, 2010:** Nonparametric Model Based Estimation in Survey Sampling
40. **Garcia, MS, 2010,** Detection of Influential Observations in Regression Model with High Dimensional Predictors.
41. **Guilatco, MS, 2010:** Nonparametric Estimation of the Switching Regression Model
42. **San Diego, MS, 2010:** Nonparametric Comparison of Predictive Accuracy
43. **Lemence, MS, 2009:** Nonparametric Testing of Structural Change and Spatial Heterogeneity in Spatial-Temporal Models.
44. **Balagbis, MS, 2009:** Estimation of Panel Models with Structural Change
45. **Santos L, MS, 2009:** Small Sample Estimation in Dynamic Panel Data Models
46. **Bastero, MS, 2009:** Robust Estimation of a Spatiotemporal Model in Epidemic.
47. **Campano, MS, 2008:** An Estimation Procedure of an ARIMA Model in the Presence of Structural Change.
48. **Sobrevenas, MOS, 2008:** A Nonparametric Analysis of the Impact of Rice Trade Policy Reform on the HH Welfare in the Philippines.
49. **Mina, MOS, 2008:** Unearthing Poverty with MARS: Application of Multivariate Adaptive Regression Splines in Identifying Household Poverty Correlates in the Philippines.
50. **Gatpatan, MS, 2008:** Estimation of a Response Surface Model Using Backfitting
51. **Dela Rosa, MOS, 2008:** Nonparametric Krigging Using Additive Models
52. **Martinez, MS, 2008:** An Estimation Procedure for Multivariate Spatial Temporal Model in Small Area Estimation
53. **Dumanjug, MS, 2007:** Bootstrap Procedures in an Iteratively Estimated Spatial-Temporal Model Typhimurium in Citrus Systems
54. **Lansangan, MS, 2006:** A Method of Attaining Sparsity of Principal Components in Time Series Data.
55. **Suratos, MOS, 2005**: Sparse Spatial Autoregressive Modeling of Poverty in the Philippines
56. **Komoto, MOS, 2005:** Sparse Principal Component Analysis of Sustainable Development Index
57. **Landagan, MS, 2005:** Estimation in a Generalized Spatio-Temporal Modeling
58. **Guarte, MS, 2004:** Estimation Under Purposive Sampling
59. **Medina, MS, 2002:** Computing Aspects of Markov Switching Models with Time-Varying Transition Probabilities
60. **Fermin, MS, 2001:** Fitting Logistic and Autologistic Models with Autocorrelated Errors Using Backfitting
61. **Janagap, MOS, 1996** : Estimating Genetic Parameters Using Variance Components and Simulated Data.

Recent Presentation in Conferences

1. **2015, Singapore:** Organized a Session-Robust Methods in Big Data Analysis, **9th Conference of the Asian Regional Section of the IASC, International Association of Statistical Computing-International Statistical Institute.**
2. **2015, Rio de Janeiro, Brazil:** Robust Estimation of a Multilevel Model with Structural Change, **E. Barrios**, MJ Esmenda, **60th World Statistics Congress of the International Statistical Institute.**
3. **2015, Rio de Janeiro, Brazil:** Simultaneous Dimension Reduction and Prediction Optimization: Methods and Applications to High Dimensional Data ,**J. Lansangan**, **E. Barrios**, **60th World Statistics Congress of the International Statistical Institute.**
4. **2014, Kuala Lumpur, Malaysia:** Organized a Session– Computational Statistics, **Regional Statistics Conference, International Statistical Institute.**
5. **2014, Geneva, Switzerland:** Nonparametric estimation of a switching regression model , **E. Barrios**, R. Guilatco, **21st International Conference on Computational Statistics and 5th IASC World Conference.**
6. **2014, Geneva, Switzerland:** Semiparametric principal components Poisson regression on clustered data , K. Manalaysay, **E. Barrios**, **21st International Conference on Computational Statistics and 5th IASC World Conference.**
7. **2014, Geneva, Switzerland:** Estimation of isotonic spatio-temporal model with clustering , **M. Lucagbo, E. Barrios**, **21st International Conference on Computational Statistics and 5th IASC World Conference.**
8. **2014, Geneva, Switzerland:** Simultaneous dimension reduction and variable selection in modeling high dimensional data, **J. Lansangan, E. Barrios**, **21st International Conference on Computational Statistics and 5th IASC World Conference.**
9. **2014, Geneva, Switzerland:** Nonparametric hypothesis testing for clustered survival model , **J. Eustaquio, E. Barrios**, **21st International Conference on Computational Statistics and 5th IASC World Conference.**
10. **2014, Maryland, USA:** Computational Statistics Methods in Spatial-Temporal Models, **E. Barrios**, *Invited Paper,* **8th Annual Statistics and Probability Day, University of Maryland Baltimore County**
11. **2013, London, UK:** A Control Chart Design for Monitoring Autocorrelated Processes with Multiple Exogenous Inputs Under Model Uncertainty, S. Poblador, **E. Barrios**, **6th International Conference of the ERCIM WG On Computational and Methodological Statistics.**
12. **2013, Hongkong, SAR:** Semiparametric Poisson Regression Models in Clustered Data, E. De Vera, **E. Barrios**, **59th World Statistics Congress of the International Statistical Institute.**
13. **2013, Hongkong, SAR:** Using principal component scores as stratification variable: An alternative to multiple frame sampling methodology, **K. Santos, E. Barrios**, **59th World Statistics Congress of the International Statistical Institute.**
14. **2013, Hongkong, SAR:** Modelling clustered survival data with cured fraction, **A. Nalica, I. Gauran, E. Barrios**, **59th World Statistics Congress of the International Statistical Institute.**
15. **2013, Hongkong, SAR:** Dynamics of development in rural communities, **E. Barrios**, **G. Sarte**, **59th World Statistics Congress of the International Statistical Institute.**