የኢትዮጵያ ስታቲስቲክስ ባለሙያዎች ማኅበር

Ethiopian Statistical Association (ESA)

The 31th Annual Conference of Ethiopian Statistical Association

"Implementation of Statistical Methodology in Health and Agriculture in Ethiopia"

Sponsors

ETHIOPIAN STATISTICS SERVICE

UNICEF

UNECA-ACS

01 – 02 July, 2023

Addis Ababa, Ethiopia

Conference Program and Book of Abstracts

Venue: - Melka Addis International Hotel

Program

ESA 2023

The 31th Annual Conference of Ethiopian Statistical Association (ESA)

01 – 02 July, 2023, Addis Ababa, Ethiopia

Venue: - Melka Addis International Hotel

Sponsored by:
Ethiopian Statistics Service
UNECA-ACS
UNICEF

Addis Ababa, Ethiopia

LIST OF ABSTRACT REVIEWERS

1. Awoke Seyoum, PhD
2. Bedilu Alamirew, PhD
3. Cheru Atsmegiorgies, PhD
4. Dejen Tesfaw, PhD
5. Denekew Bitew, PhD
6. Derebachew Asfaw, PhD
7. Derejie Denbe, PhD
8. Girma Taye, PhD
9. Habte Tadesse, PhD
10. Haile Mekonnon, PhD
11. Jemal Ayalew, PhD
12. Kasahun Takele, PhD
13. Seifemichael Amare, PhD
14. Selamawit Serka, PhD
15. Taddesse Kassahun, PhD
16. Teshager Zerihun, PhD
17. Tsedeke Lamnore, PhD
18. Zeytu Gashaw, PhD

CONFRENCE ORGANIZING COMMITTEES

1. Scientific and Publication Committee

Zeytu Gashaw (PhD), Chair

Genanew Timerga, Assistant Prof.

Dejen Tesfaw, PhD

Tegodie Hibstu

Samuel Hailu

2. Logistics and Finance Committee

Taddesse Kassahun (PhD), Chair

Gezahegn Getahun

Dejen Tesfaw, PhD

Haymanote Zeleke

Samuel Hailu

Objectives of the Conference

The main objectives of the conference are to:

- Promote proper utilization of statistical methods in survey design, data collection and analysis;
- Signify applications of statistics in the areas of agriculture and health;
- Establish network among professionals working in various offices;
- Create a medium for practitioners and young statisticians to gain insights through paper presentations, lectures, and panel discussions;
- Share best practices of statistical applications in agriculture and health, and identify gaps that need improvements.

Expected outcomes of the conference

The following outputs are expected upon conducting the 31st Annual Conference of ESA:

- Bring professionals working on agriculture and health on the same page towards effectively employing statistics for the success of the Government's ten-year plan.
- Create and strengthen network of statistics professionals working in various offices.
- Open gateways for different offices to boost quality data generation and utilization.
- Find out open problems for research that are emanated from paper presentations, panel discussions and dialogues.
- Strengthen partnership of ESA with development partners

General Information

Conference Location

Melka Addis International Hotel around Piassa Churchil Ave, Addis Ababa, Ethiopia

Registration

The registration desk at the **Melka Addis International Hotel** will be open from 08:00 AM on Saturday 01st July 2023.

Internet Access

Free wireless Internet access will be provided to ESA 2023 conference participants during the days of the Conference at the Hall. The information needed in order to access the wireless network will be provided at the registration desk at the conference venue.

Liability

The Conference fees DO NOT include provisions for the insurance of participants against personal injuries, sickness, and theft or property damage. This also applies to any event held during the Conference period. Neither the Conference Organizing Committee nor its sponsors nor committee members assume any responsibility for loss, injury or damage to persons or belongings, however, caused.

CONFERENCE PROGRAM

The 31st Annual Conference of ESA, Ethiopia

Addis Ababa, Ethiopia,

01 –02, July 2023

Venue: Melka Addis International Hotel

SATURDAY, 01st July 2023	
08:00-9:00	**Registration**
SESSION 1: OPENING SESSION	
9:00-09:30	**Master of Ceremony**: **Ms. Aberash Tariku**, Ethiopian Statistics Service **Welcoming Speech: Dr. Dejen Tesfaw**, President of ESA **Official Opening**: Guest of honor H.E **Dr. Fitsum Assefa**, Minister, Ministry of Planning and Development **Keynote Speech 1: Dr. Beker Shale**, Director General, Ethiopian Statistics Service **Keynote Speech 2: Oliver Chinganya**, Director General, Center of Statistics for Africa **Keynote Speech 3: XXXX, Country director of** UNICEF
SESSION 2: INVITED TALKS 1	
Chair: Mr. Biratu Yigezu, Advisory, Ministry of planning and development and **Dr. Emanuel Gabreyohannes**, Ethiopian Civil Service University	
09:30-10:45	1. **Dr. Girma Taye**, Addis Ababa Univirsty • Health and or Agriculture 2. **XXXXXXX,** Ethiopian Statistical Service • 3. XXXXXXX UNICEF • 4. Mr. Anteneh Tesema,TASK Clinical Research Institute • Statistical Principles for Clinical Trials
10:45-11:00	**Health Break**

11:00-12:30	**Discussion**
12:30-14:00	**Lunch Break**
SESSION 3: Parallel Session: Applied Statistics	**Room 1**
Chair: Dr. Tsedeke Lamnore, Wachamo University **Rapporteur**: Dr. Bedilu Alamirie,, Addis Ababa University	
14:00 - 15:30	*15 minutes for each presenter* **Paper 1:** Hierarchical Bayesian Spatial Small Area Model for Binary Data Under Spatial Misalignment **Speaker:** Kindie Fentahun **Paper 2:** Genotype selection for Grain yield of Sorghum through Generalized Linear Mixed Model **Speaker:** Mulugeta Tesfa **Paper 3:** Modeling oil seed export potential in Ethiopia: Inferences from a dynamic gravity approach **Speaker**: Lencho Idiris **Discussion**
SESSION 3: Parallel Session: Bio-Statistics	**Room 2**
Chair: Dr. Shiberu Temessgen, Addis Ababa University **Rapporteur:** Dr. Awoke Seyoum, Bahir Dar University	
14:00 – 15:30	*15 minutes for each presenter* **Paper 1. Joint Modeling of Longitudinal Pulse Rate and Time-to-Default from Treatment of Congestive Heart Failure Patients** in Felege-Hiwot Referral Hospital, Bahir Dar, Ethiopia **Speaker**: Yikeber Abebaw **Paper 2:** Spatio-temporal prediction of malaria incidence using supervised machine learning approaches **Speaker:** Teshager Zerihun **Paper 3:** Improving the Precisions of Survey Based Estimates of Stunting using A Small Area Estimation Approach **Speaker**: Seyifemickael Amare

	Discussion
SESSION 3: Parallel Session: Theoretical/Methods	**Room 3**
Chair: Dr. Denekew Bitew Bahir Dar University **Rapporteur**: Dr. Derejie Denbe Hawassa University	
14:00-15:30	*15 minutes for each presenter* **Paper 1:** A New Generalized-X Family of Distributions: Applications, Characterization and a Mixture of Random Effect Models Applied to Diabetes Mellitus Patients Data **Speaker:** Getachew Tekle **Paper 2:** Bayesian methods for borrowing historical information for count data **Speaker**: Akalu Banbeta **Paper 3:** Classification of Imbalanced Data using Machine Learning Algorithms to Predict the Risk of Renal Graft Failures in Ethiopia **Speaker**: Getahun Mulugeta **Discussion**
15:30-15:50	**Health Break**
SESSION 4:	**PLENARY SESSION Main Hall**
Chair: Dr. Selamawit Serka, Hawassa University **Rapporteur:** Dr. Jemal Ayalew, Wollo University	
15:50-17:00	*20 minutes for each presenter* Virtual 1 and Virtual 2 **Discussion**
SUNDAY, 02nd July 2023	
SESSION 5: PLENARY SESSION Main Hall	
Master of Ceremony: Dr. Zeytu Gashaw, ESA Addis Ababa University	
Chair: Dr. Essaye Kebede, Vice President of Bahir Dar University **Rapporteur:** Dr. Haile Mekonnon, Bahir Dar University	

9:00-9:40	1. **Prof. Solomon Harrar,** University of Kentucky, USA • Nonparametric Finite Mixtures for Overcoming Biomarker-Error Bias **Speaker**: XXXXXXXXX 2. **Prof. Josaph Beyene,** McMaster University, Canada **Discussion**
9:40-10:15	**Discussion**
10:15-10:30	**Health break**
SESSION 5: BUSINESS SESSION Main Hall	
Chair: Dr. Butte Gotu and Mr. Mekonnon Tadesse, Addis Ababa University **Rapporteur:** Seifu Neda, Haramaya University	
10:30-12:45	• ESA-EC annual work report • External audit report • Internal audit report • Annual plan for 2023/24 • Discussion on the reports • Election of EC • Closing Speech
12:45	**Lunch**
End of conference	

BOOK OF ABSTRACTS

Hierarchical Bayesian Spatial Small Area Model for Binary Data Under Spatial Misalignment

Kindie Fentahun Muchie*[1], Anthony Kibira Wanjoya[2], and Samuel Musili Mwalili[2]

1Pan African University Institute for Basic Sciences, Technology and Innovation, Nairobi, Kenya

1Bahir Dar University, Bahir Dar, Ethiopia

2Jomo Kenyatta University of Sciences, Technology and Innovation, Nairobi, Kenya

Abstract

Small area model has become a popular method for producing reliable estimates for small areas. Small area modeling may be carried out via model assisted approaches within the design-based paradigm or model-based approaches. A model assisted design-based inference may be reliable in situations when there are large or medium samples in areas, while if data are sparse, model-based approach may be a necessity. Model based Bayesian analysis methods are becoming popular for their ability to combine information from several sources as well as taking account of uncertainties in the analysis and spatial prediction of spatial data. However, things become more complex when the geographic boundaries of interest are misaligned. Some authors have addressed the problem of misalignment under hierarchical Bayesian approach. In this study, we developed and assessed the performance of non-trivial extension of existing hierarchical Bayesian model for binary data under spatial misalignment. In this study, we developed a spatial hierarchical Bayesian small area model for a binary response variable under spatial misalignment. The developed model is fusion model, considering both areal level and unit level latent processes. The process models generated from the predictors were used to construct the basis so as to alleviate the well-known problem of collinearity between the true predictor variables and the spatial random process. A simulation study demonstrated that the model has good performance.

Keywords: Small Area Estimation; Hierarchical Bayesian; Spatial Misalignment; Fusion Process

***correspondence** mkindief@gmail.com

Genotype selection for Grain yield of Sorghum through Generalized Linear Mixed Model

Mulugeta Tesfa[1, 2]*, Temesgen Zewotir[3], Solomon Assefa Derese[4], Denekew Bitew Belay[1] and Mark Laing[5]

1 Department of Statistics, College of Science, Bahir Dar University, P. O. Box, 79, Bahir Dar, Ethiopia;

2 Department of Statistics, College of Natural and Computational Sciences, Wollo University, P. O. Box, 1145, Dessie, Ethiopia;

3 Schools of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Durban, 4041, South Africa;

4 Department of Plant breeding, College of Agriculture, Woldia University, P. O. Box, 53, Woldia, Ethiopia;

5 School of Agricultural, Earth and Environmental Sciences, College of Agriculture, Engineering & Science, Pietermaritzburg, University of KwaZulu-Natal, Durban, 4041, South Africa;

Abstract

The use of the classical model provides a correct analysis only if all the effects are fixed. For experiments that include both fixed and random effect, general linear mixed model is one of the appropriate models to handle the non-normal distributed response variables. The aim of this study is to perform the genotype selection through generalized linear mixed model and identify the impact of treatment and the related traits on the grain yield. The data were collected using lattice square design and measured the phenotype traits of sorghum. The result of PCA was used as an input variable for the general linear mixed model. The data analysis was performed using a general linear mixed model with maximum likelihood methods to estimate the parameters of the model. The result showed that the grain yield had a Gamma distribution and there was a treatment effect on grain yield. The first principal component was significant for grain yield. The variability of grain yield due to the random effects, that were replication within treatment, genotype and the interaction of genotype by treatment, were significant. The best genotypes, that are effective for mass production of sorghum, were G137, G66 and G156 under stress conditions and G55, G41 and G78 under irrigated conditions. Overall, genotype selection using a general linear mixed model for grain yield is recommended for genotype selection of plant breeding.

Keywords: General Linear Mixed Model; Genotype performance; Random effect; non-normality

Modeling oil seed export potential in Ethiopia: Inferences from a dynamic gravity approach

Lencho Idiris[1*] and Aboma Temesgen [1]

[1]Department of Statistics, College of Computing and Informatics, Haramaya University, P.O. box138 Dire Dawa, Ethiopia

Abstract

Ethiopia's oil seed export provides income to millions of growers and others market actors along the value chain and it is the second among commodities generating currency for the country. Though there were plentiful empirical evidence in modeling oil export potential in Ethiopia, communalities in terms of econometric techniques is disperse. This inconsistence poses a problem to make dependable and consistent inference across economies. Thus, this study aimed to model oil seed exports potential (in the dimension of export sales) in Ethiopia's via more realistic model application, dynamic panel gravity model. The study utilized panel data that comprising 46 countries of the Ethiopia's oil seed importers for 19 years from the period 2002 to 2020. The model was estimated with system GMM estimator. The estimated model result revealed that 1% increase in lagged EOSE, real GDP of Ethiopia, population of importing countries, openness to trade of both Ethiopia and importing countries, and foreign direct investment inflow of Ethiopia, and weighted distance, ceteris paribus, turn out were found to increase Ethiopia's oil seed exports potential by around 0.207%, 1.98%, 0.424%, 1.93%, 1.02%, 0.188%, and decrease by 1.314%, respectively. Moreover, the magnitude of oil seed export potential was found the highest with Asian, European and then African countries as a continent. The analyses also implied policies that would promote foreign direct investment, supply capacity, trade liberalization, and relatively cheaper transportation costs in order to progress the Ethiopia's oil seed exports performance. This study recommended that export diversification and bilateral trade negotiations stepping up the current status of the external sector are the necessary steps to exploit Ethiopia's untapped oil seed export potential.

Keywords: Export potential, Oil seed, Generalized Methods of Moment, Dynamic Gravity Model

***Corresponding author: E-mail:** lenchoidris18@gmail.com

Joint Modeling of Longitudinal Pulse Rate and Time-to-Default from Treatment of Congestive Heart Failure Patients in Felege - Hiwot Referral Hospital, Bahir Dar, Ethiopia

Yikeber Abebaw[1*]; Kasim Mohammed[2]; Adem Aragaw[3]; Bezanesh Melese[1]
1Department of Statistics, Debre Tabor University, Debre Tabor, Ethiopia;
2 Department of Statistics, University of Gondar, Gondar, Ethiopia;
3Department of Statistics, Mizan Tepi University, Mizan Tepi, Ethiopia

Abstract

Background: Globally, heart failure is a rapidly growing public health problem with an estimated prevalence of >37.7 million. It is a shared chronic phase of cardiac functional loss secondary to many etiologies. The main purpose of this study was to investigate the risk of longitudinal change in pulse rate on time to default from treatment among congestive heart failure patients. Hospital-based retrospective studies were conducted among 302 congestive heart failure patients who were 15 years old or older and who were on treatment follow-up from the 1 February 2016 to 31 December 2018 in Felege - Hiwot Referral Hospital, Bahir Dar, Ethiopia. Data were analyzed using SAS and R software. First, data were analyzed using linear mixed model and survival models separately, and then the joint models of both sub-models were analyzed using joint model analysis. Out of the total 302 respondents, 34.1% of the respondents defaulted from treatment. About 55.2% of male respondents are defaulting and the remaining was censored. The results for separate and joint models were quite similar to each other but not identical. However, the estimated association parameter (α) in the joint model is (HR = 1.0311, 95% CI: 1.0033, 1.0597, P = 0.0278), providing there is evidence of a positive significant relationship between the survival and the longitudinal sub-models. Thus, defaulting is more likely to occur in patients with higher pulse rates. A patient, who is male, New York Heart Association class IV, had low left ventricular ejection fraction and comorbid with hypertensive, chronic kidney disease, pneumonia were risk factors of pulse rate change and defaulting from the treatment of congestive heart failure patients. The joint model was preferred for simultaneous analyses of repeated measurement and survival data. Thus, the longitudinal measure pulse rate had a positive significant effect on time to default from the treatment of patients.

Keywords: Pulse Rate, Congestive Heart Failure, Time to Default, Survival Analysis, Longitudinal Analysis, Joint Model

Spatio-temporal prediction of malaria incidence using supervised machine learning approaches

Teshager Zerihun Nigussie[1*], Temesgen Zewotir[2] and Essey Kebede Muluneh[3]

[1]Department of Statistics, College of Natural and Computational Sciences, Debre Tabor University, Debre Tabor-Ethiopia

[2]School of Mathematics, Statistics and Computer Science, College of Agriculture Engineering and Science, University of KwaZulu-Natal, Durban, South Africa

[3]School of public health, College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar-Ethiopia

Abstract

Malaria is a leading cause of morbidity that disrupts the agricultural activities of societies in developing nations. The two-decade decreasing trajectory of malaria cases has been disrupted globally since 2018. Hence, prediction of malaria incidence attracted the attention of health care practitioners to prioritize malaria prevention and elimination activities in the malarious regions. This study is aimed to predict malaria incidence using supervised machine learning approaches. The monthly malaria incidence of districts obtained from the Amhara Public Health Institute. Environmental data were obtained from the NOAA, CHIRPS, and Woldclimate. Monthly malaria incidence is predicted using classification and regression tree for spatiotemporal data (CAST) and Bayesian spatiotemporal models. The predictive accuracy of models was evaluated using validation statistics. The models were trained and tested using 80/20 ratio of training/testing dataset divisions and 10-fold cross-validation strategies. The altitude of districts, year, season, one-month lag of NDVI and soil moisture, and two-month lag of relative humidity had a significant effect on malaria incidence. Among these, altitude of districts and seasonal variation were selected and offered an optimistic prediction of malaria incidence in the region. The CAST and Bayesian spatiotemporal model with AR(2) random effect had a smaller validation statistics for prediction. The AR(2) random effect model has been used for prediction via a contagious spatial block (LLO) and LTO training/testing data partitions. The LTO data partition offered a prediction with smaller WAIC but a beat higher RMSE and MAE than LLO. The CAST with a target-oriented 10-fold-cross validation (LLO) also gave optimistic malaria incidence predicted values using seasonality, spatial coordinates and average altitude as predictor variables. The CAST and Bayesian spatiotemporal models were supervised machine learning techniques that can be used to predict monthly malaria incidence for elimination and control of malaria in the region.

Keywords: Prediction, Data division, Cross-validation, Predictive accuracy, Bayesian spatio-temporal, Malaria

Improving the Precisions of Survey Based Estimates of Stunting using A Small Area Estimation Approach

Seyifemickael Amare Yilema[1*], Yegnanew A Shiferaw[2]

[1]Department of Statistics, College of Natural and Computational Science, Debre Tabor University, Debre Tabor, Ethiopia; samarey1981@gmail.com

[2]Department of Statistics, University of Johannesburg, Auckland Park Kingsway Campus,

P.O. Box 524, Johannesburg 2006, South Africa; yegnanews@uj.ac.za

[*]Corresponding author: samarey1981@gmail.com

Abstract

A survey is often designed to produce reliable estimates of target variables of interest at national and regional levels. However, the survey is not available below the reginal level administration such as zones. But reliable estimates are needed in many ways for zones, which are unplanned domains (areas). Disaggregated direct survey estimates for zones are unreliable due to small sample sizes. This paper aims to improve the precision of direct survey estimates of stunting at zonal levels using small area estimation (SAE) under the Fay-Herriot (FH) model. More precise and reliable estimated are produced for small sample sizes of unplanned domain (areas) using SAE under FH model. However, unprecise and unreliable estimates are produced for large samples sizes such as Addis Ababa, Dire Dawa and Harreri. This is clearly shows that SAE approaches are important for areas with small sample sizes. For almost all zones in Ethiopia, the proposed empirical best linear unbiased predictor (EBLUP) estimates outperform survey-based estimates in terms of root mean squared error (MSE) and coefficient of variation (CV). Indeed, information on zonal level stunting can help governments for policy formulations and planners.

Key words: Stunting, small area estimation, census, survey, linking, precision

A New Generalized-X Family of Distributions: Applications, Characterization and a Mixture of Random Effect Models Applied to Diabetes Mellitus Patients Data

Getachew Tekle[1]*, Rasool Roozegar[2], G. G. Hamedani[3]

1 Department of Statistics, Wachemo University, Hossana, Ethiopia, getch.55tekle@gmailcom

2Department of Statistics, Yazd University, Yazd, Iran, rroozegar@yazd.ac.ir

3Department of Mathematics, Statistics and Computer Science, Marquette University USA, gholamhoss.hamedani@marquette.edu

Abstract

The researchers in applied statistics are recently highly motivated to introduce new generalizations of distributions due to the limitations of the classical univariate distributions. In this study, we propose a new family called new generalized-X family of distributions. A special sub-model called new generalized-Weibull distribution is studied in detail. Some basic statistical properties are discussed in depth. The performance of the new proposed model is assessed graphically and numerically. It is compared with the five well-known competing models. The proposed model is the best in its performance based on the model adequacy and discrimination techniques. The analysis is done for the real data and the maximum likelihood estimation technique is used for the estimation of the model parameters. Furthermore, a simulation study is conducted to evaluate the performance of the maximum likelihood estimators. Additionally, we discuss a mixture of random effect models which are capable of dealing with the overdispersion and correlation in the data. The models are compared for their best fit of the data with these important features. The graphical and model comparison methods implied a good improvement in the combined model.

Keywords: NG-X Family; NG-Weibull distribution; Weibull distribution; Simulation; Combined model; Random effect models; Overdispersion; Correlation.

Bayesian methods for borrowing historical information for count data

Akalu Banbeta[1, 2,*], Emmanuel Lesaffre[3], Reynaldo Martina [4] and Joost van Rosmalen [5, 6]

I-Biostat, UHasselt, Hasselt, Belgium, [2]Department of Statistics, Jimma University, Jimma, Ethiopia

[3]I-Biostat, KU Leuven, Leuven, Belgium

[4]Faculty of Science, Technology, Engineering and Mathematics, Open University, Milton Keynes, UK

[5]Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands

[6]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

Abstract

Including historical data into the analysis of a current clinical trial may reduce the necessary sample size and/or increase the power of analysis. However, this only applies when the historical data are similar enough to the current data. Several Bayesian approaches for borrowing historical information in a dynamic way have been proposed, such as the meta-analytic-predictive (MAP) prior and the modified power prior (MPP) both for a single historical study as well as for multiple historical studies. Here, we examine the performance of the MAP and MPP approaches for the analysis of (overdispersed) count data when multiple historical control data are incorporated into the analysis of the current data. To this end, we explore the Poisson and the negative binomial distribution. We propose a computational approach based on path sampling algorithm for the MPP approach. We illustrate our approach using the data of a RCT involving patients with an overactive bladder whereby the response is the frequency of incontinence periods. Further, we have conducted a simulation study in case of heterogeneity of the control arms across trials. For similar current and historical control arms, the MPP approach offers greater statistical power than the MAP approach. In this case, the MPP approach could help to reduce about one-third of the required sample size in a future study by borrowing historical information. When the means are different across the control arms, the MPP approach yields a slightly inflated type I error rate, whereas the MAP does not. When the dispersion parameters are different across the control arms, the results are reversed. In conclusion, the MPP approach outperforms the MAP approach for the count data in terms of power of test and type I error rate.

Keywords: Count Data, meta-analytic predictive prior, negative binomial, Poisson, over-dispersion, power prior

* Presenting author, Email address: akalubanbeta.tereda@student.uhasselt.be

Classification of Imbalanced Data using Machine Learning Algorithms to Predict the Risk of Renal Graft Failures in Ethiopia

[1*]Getahun, M., [2]Temesgen, Z., [1]Awoke, S., [3]Mahteme, B., & [3]Leja, H.

[1]Department of Statistics, Bahir Dar University, Ethiopia
[2]School of Mathematics, Statistics, and Computer Science, KwaZulu Natal University, South Africa
[3]St.Paul's Hospital Millennium Medical College, Ethiopia

Abstract

The prevalence of end-stage renal disease has raised the need for renal replacement therapy over recent decades. Even though a kidney transplant offers an improved quality of life and lower cost of care than dialysis, graft failure is possible after transplantation. Hence, this study aimed to predict the risk of graft failure among post-transplant recipients in Ethiopia using the selected machine learning prediction models. The data was extracted from the retrospective cohort of kidney transplant recipients at the Ethiopian National Kidney Transplantation Center from September 2015 to February 2022. In response to the imbalanced nature of the data, we performed parameter tuning, threshold moving, model stacking ensemble, and probability calibrations to improve the prediction results. Merit-based selected probabilistic (logistic regression, naive Bayes, and artificial neural network) and tree-based ensemble (random forest, bagged tree, and stochastic gradient boosting) models were applied. Model comparison was performed in terms of discrimination and calibration performance. A total of 278 completed cases were analyzed, with 21 graft failures and 3 number of event per predictor. From the comparison of models at the individual level, the bagged tree and random forest have top and equal discrimination performance (AUC-ROC = 0.84). In contrast, the random forest has the best calibration performance (brier score = 0.045). Under the stacking ensemble learning, the result from stochastic gradient boosting as a meta-learner has the top discrimination (AUC-ROC = 0.87) and calibration (brier score = 0.048) performance. Regarding feature importance, chronic rejection, blood urea nitrogen, number of post-transplant admissions, phosphorus level, acute rejection, and urological complications are the best predictors of graft failure. The tree-based ensembles outperform probabilistic models in terms of discrimination and calibration due to the help of probability calibration. Stacking ensemble learning is a good boosting of prediction performance for imbalanced data.

Keywords: Renal Transplantation; Graft Failure; Imbalanced Data; Tree-Based Ensembles, Stacking Ensemble; Probabilistic Models

Ethiopian Statistical Association Action Plan 2022/23

Five year strategic plan (2023 – 2027)

Part I: Research and Project Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Project implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Strategic Issue One : Undertake project need assessment																		
1.1	Undertake project need assessment on UN Agencies , International Organizations , INGOs and research institutes	No	20														Project Coordinator ESA-EC	5,000
1.2	Signing MOU with development partners	No	3	Signed MOU														
Strategic Issue Two: Proposals Development and Documentary Research																		
2.1	Writing an innovative and sound grant project /or research proposals	No	2	Developed innovative proposals													Project Coordinator, ESA-EC	Based on rate
2.2	Develop a wining project proposals for call for project proposals	No	2	Developed proposals													Project Coordinator, ESA-EC	Based on rate
2.3	Conduct Documentary Research on ESA's past, present and Future perspective	No	1	Produced documentary research													Project Coordinator	50,000
Strategic Issue Three: Capacity building training (Business Development) and Award																		
3.1	Organize and provide an intensive project management training for stakeholders with allocated budget (If the project awarded to ESA)	Time	1	Delivered training													Project and Training Coordinators and ESA-EC	100,000
3.2	Organize and provide an innovative grant project proposal writing for	Time	1	Delivered training													Project and Training	100,000

	stakeholders with budget (If the project awarded to ESA)																Coordinators and ESA-EC	
3.3	Organize and provide a result based Project Monitoring and Evaluation training for stakeholders with allocated budget (If the project awarded to ESA)	Time	1	Delivered training													Project and Training Coordinators and ESA-EC	100,000
3.4	Provide award for best project and research proposal writers for successful projects/researches	No	5	Award													Project Coordinators and ESA-EC	100,000
3.5	Organize free Virtual Advanced Project Management and Project M&E training for EC members and members	No	100														Project Coordinators and ESA-EC	20,000
Total Project Cost																		

Part II: Training Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Training implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
1.	**Identifying potential stakeholders and distribute training collaboration letters**																	
1.1	Identify potential stakeholders	No	25	Identified stakeholders													Training Coordinator	
1.2	Conduct training need assessment	No	25	Conducted assessment													Training Coordinator	6,000
1.3	Develop training collaboration letters and Disseminate it	No	25	Developed collaboration letters													Training Coordinator	
1.4	Signing MoU	No	5	Signed MOU													ESA President	2,500
2.	**Develop Training proposals**																	

2.1	Developing training proposals		No	12	Develop proposals													Training Coordinator, ESA-EC,	72,000
3.	**Develop Database of Trainers**																		
3.1	Vacancy Announcement for trainers (per training)		No	1	Set requirements													Training Coordinator	
3.2	Assign Ad-hoc committee to select potential trainers		No	1	Set selection criteria													Training Coordinator, ESA-	
3.3	Design Database for selected trainers		No	1	Design database for list trainers														
4.	**Fix Training Date and Announce Call for Trainings**																		
4.1	Assign trainer from the database		Time	12	Assign trainer													Training and Project Coordinators and ESA-EC	
5.	**Trainings for individually registered trainees**																		
5.1	For Individual	Stata	No	40														Training & Project Coordinators	70,000
5.2		SPSS	No	40														Training & Project Coordinators	70,000
5.3		R	No	60														Training & Project Coordinators	90,000
5.4		ODK	No	50														Training & Project Coordinators	70,000
5.5		Survey Solution	No	30														Training & Project Coordinators	45,000
6.	**Offer Free Trainings for Members and Invited Guests**																		
6.1	Ethiopian Statistics Service	Advanced data Management and Analysis using SPSS/STATA/ R	No	60														Training & Project Coordinators	90,000

6.2	Financial sectors Banks & Insurances	Marketing research and analysis	No	200														Training & Project Coordinators	360,000
		Advanced time series & panel data analysis using STATA/ R	No	130														Training & Project Coordinators	225,000
6.3	Ministry of National Planning & Deve't	Advanced Research Methodology	No	30														Training & Project Coordinators	45,000
		Advanced data Management and Analysis using SPSS/STATA/ R	No	30														Training & Project Coordinators	45,000
6.4	Ministry of Agriculture	Advanced research methods and analysis using SPSS/ SAS/STATA/R	No	50														Training & Project Coordinators	90,000
6.5	Ministry of Finance	Advanced research and data Analysis using SPSS/STATA/R	No	30														Training & Project Coordinators	45,000

Part III: Chapter Offices and Members' Affairs Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Project implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Strategic Issue One : Design Online Membership Services																		
1.1	Establish online membership registration	Time	1	Design online registration platform													Chapter Offices and Members Coordinator ESA-EC	
1.2	Establish online membership payment	No	1	Design online registration platform													Chapter Offices Coordinator	

Strategic Issue Two: Evaluate Current Individual and Institutional Members																			
2.1	Identify active and inactive members	No	1	Evaluate members status													Chapter Offices	
2.2	Design interactive web based database for Members	No	1	Designed database													Chapter Offices and Members Coordinator ESA-EC	
2.3	Increase the number of individual member	No	1000														Chapter Offices and Members Coordinator ESA-EC	
2.4	Increase the number of Institutional members	No	10														Chapter Offices and Members Coordinator	
2.5	Undertake membership satisfaction survey	Time	1															
Strategic Issue Three: Evaluate and Capacitate Chapter Offices																			
3.1	Evaluate the existing status and activities of chapter offices	No	1	Examine status of chapter offices													Chapter Offices and Members Coordinator	
3.2	Assist and capacitate Chapter offices	Time	12														Chapter Offices and Members Coordinator	
3.3	Launch additional Chapter Offices	No	5														Chapter Offices and Members Coordinator	
Strategic Issue Four: Offer Free Training at Chapter Offices																			
4.1	Offer free trainings at Chapter office	Time	6	Provide trainings													Training and Chapter Offices	ESA will pay only Trainers fee

Total Project Cost																			

Part IV: Secretariat Office Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Project implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Strategic Issue One : Organize Meeting, Annual Conference, National Workshop																		
1.1	Organize monthly ESA-EC meeting	Time	12	Organized meeting													ESA-EC	175,00
1.2	Organize National Workshop on the role of Statistician for national economic growth and development	No	1	Organize workshop													ESA-EC	75,000
1.3	Organize 31th upcoming Annual Conference	Time	1	Work on conference works													ESA-EC	TBF
1.4	Organize joint AB-EC Meeting	No	3	Conduct meeting													ESA-EC	TBF
1.5	Organize Stakeholders Meeting	Time		Conduct meeting													ESA-EC	TBF
Strategic Issue Two : Develop ESA' Strategic Plan																		
2.1	Develop ESA's Strategic Plan	No	1	Developed SP													ESA-EC	50,000
Strategic Issue Three: Managing the implementation of JESA																		
3.1	Work on JESA continuity	Time	3	Examine status of chapter offices													ESA-EC, Chief editor	

3.2	Work on JESA to be indexed	Time	12	Make it indexed													ESA-EC, Chief editor	
Strategic Issue Four: ESA Office Building Project																		
4.1	Assign Ad-hoc Committee	No	1	Assign Committee													ESA-EC	
4.2	Organize Consultative Meeting	No	2														ESA-EC and AB	
4.3	Organize income generating projects	No	2														ESA-EC	
Strategic Issue Five: Preparing Performance Report																		
5.1	Preparing Administrative Performance Report	No	4	Prepare performance report													Admin	
5.2	Follow-up the implementation of ESA-EC decisions	Time	12	Assess EC decisions													Secretariat	
Total Cost																		

Part V: Finance Administration Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Project implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Strategic Issue 1: Financial Monitoring, Reporting and Documentation																		
1.1	Monitoring ESA's revenues, assets, expenses and financial transactions	Time	12	Financial Documentation													Finance	3,600

1.2	Prepare financial statement on quarterly basis	Time	4	Produced financial statement													Finance	
1.3	Prepare annual financial report	Time	1	Produced financial report													Finance	
1.4	Increase ESA's revenue by 50% from 2022 to 2023	Time	12	Increased revenue													Finance and EC members	
1.5	Reduce administrative expenses by 10% from 2022 to 2023	Tim e	12	Reduced administrative expenses													Finance and EC members	
Strategic Issue 2: Financial Capacity Building																		
2.1	Organize Financial Management training for EC members and regular staff	Time	1	Organized training													Finance and Training Coordinators and ESA-EC members	25,000

Part VI: Audit and Inspection Work Plan

S.N	Key components and activities	Unit	Target	Milestone	Project implementation period												Responsible	Estimated budget in ETB
					June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May		
Strategic Issue One : Audit Control																		
1.1	Follow-up Audit	Time/ month	12														Internal Auditor	3,600
1.2	ESA audit assistance (professional help)	Time	12														Internal Auditor	
1.3	Audit the 2021/2022 ESA account	No	6														Internal Auditor	

1.4	External audit examination of activities of 2021/22	No	6														Internal Auditor	
1.5	Internal audit service (Internal Control)	No	6														Internal Auditor	
1.6	ESA Revenue and expenditure monitoring	No	6														Internal Auditor	
1.7	General work and budget control of ESA	Time/ month	12														Internal Auditor	
Strategic Issue Two : General Inspection																		
2.1	Inspect the 2021/2022 ESA Asset	No	4														Internal Auditor	
2.2	Facilitates the disposal of property on the basis of their service and age	No	1														Internal Auditor	
2.3	Monitoring and evaluating work and planning performance	No	3														Internal Auditor	
2.4	Monitoring the executive committee decisions	No	6														Internal Auditor	
2.5	Prepare ESA annual audit and inspection report	No	1														Internal Auditor	
Total Cost																		3,600