

Predicting seed yield of *Ficus* fruits by fruit dimensions

H.L.D.K. Jayarathna¹, L.S. Nawarathna^{1*} and W.A.I.P. Karunaratne²

¹Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka, ²Department of Zoology, University of Peradeniya, Sri Lanka
*lakshikas@pdn.ac.lk

Ficus is one of the largest plant genus which has an ecological significance due to the presence of “keystone” species. Availability of its sole mutualistic wasp pollinator and the effect of non-pollinator wasps determine the availability of seeds in *Ficus* fruits to produce the next generation of each species. In most of the previous studies on seed yield of *Ficus* fruits, seeds have been counted manually, which is a time consuming, hectic process. Therefore, the main objective of this study is to introduce a model for predicting seed yield in two *Ficus* species. *Ficus racemosa* and *Ficus callosa* trees which are located in Kandy municipal area, Thumpane & University of Peradeniya were selected to obtain about 120 *Ficus* fruits from each species. Fruit length and two fruit diameters as diameter1 and diameter2 were measured. Additionally, pollinators & non-pollinators reared from each fruit were counted and recorded. Local polynomial regression and generalized additive models were used for constructing these models to predict the number of seeds per fruit. Correlation analysis was carried out to identify the relationship between number of pollinators and seed yield per fruit. Seed predicting models for *Ficus racemosa* and *Ficus callosa* were validated using Mean Squared Error (MSE) of testing samples and AIC, BIC values were used for selecting the best model.

Two models which were constructed for Kandy municipal and Thumpane were best described with least MSE of testing samples and moderately large R^2 values. Fruit length was taken as a single predictor for both models. It can be concluded that for predicting seed yield there is no need of measuring two diameters since the best model requires only fruit length. Merging observations of two species for single area gave a better result rather than predicting seed yield for separate species. This implies that type of species does not play an important role to predict seed yield of *Ficus callosa* and *Ficus racemosa* using generalized additive model with fruit length. This study reveals that when seed yield is less than 1000 it gives more accurate predictions with respect to higher seed yield. Poisson regression modeling gave a better result for modeling in *Ficus callosa* with min-max scaled variables with lowest MSE value of the testing sample. By using local polynomial regression curves, it was identified that biasedness and variance both can be optimized using optimal bandwidth and it gives freedom to the flow of data by keeping non-parametric qualities.

Authors are grateful to National Science Foundation (Grant No: RG/2011/NRB/01) for funding this study.