Volume 6, Issue 3, March 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Software Model for Estimation and Prediction of Unevenly Spaced Livestock Population

[1]S. Nithiyanantha Vasagam, [2]K M Jagathnath Krishna, [3]K Giriyappa
[1, 3] Economics Research Division, CSIR- Central Leather Research Institute, Adyar, Chennai, Tamilnadu, India
[2] Division of Cancer Epidemiology and Biostatistics, Regional Cancer Centre, Medical College Campus, Thiruvananthapuram, Kerala, India

Abstract - The prediction software for livestock population and availability of hides and skins, named as LivSoft is a simple software tool for the prediction of livestock population and estimation of availability of hides and skins from the predicted livestock population. This software is developed for the benefit of leather industry/sector. Database on livestock population has been obtained from livestock census data. It automatically does the calculation by using the parameters like annual growth rate and standard error, basically which forms an input for the statistical model to predict the livestock population and thereby estimating the availability of hides and skins. The predicted livestock population is validated / compared with the regression model. This version will predict the livestock population at the national level.

Key words: Statistical Modelling, Database, Regression, Growth Rate, Net Off-Take Rate, Livsoft

I. INTRODUCTION

The leather industry in India holds an important position in contributing towards the economy of the nation as it is one of the oldest manufacturing industries. Recently, the export of leather and leather products gained great momentum and the export of Indian leather products have shown marginal growth. This is because of the best planning made by the industry for optimum utilization of available raw materials and resources. Over several years of its existence, the leather industry in the nation has undergone drastic changes from being a mere exporter of raw materials in the early 60's and 70's. Now, the country has become one of the largest exporters of value-added leather products and finished goods due to the contribution made by the big and small scale industries. In India, the leather sector is spread across and produces a comprehensive range of products from raw hides to fashionable shoes. The industry consists of firms in all capacities, including small artisans to major global players. These have been an increasing emphasis on the planned development of industry, which is aimed at optimum utilization of available raw materials for maximum returns, especially from exports.

The basic raw material for leather industry, the hides and skins are obtained as a by-product of meat industry in India. Hence the leather industry gets the raw materials from two sources, majority from slaughtering and the rest from fallen animals. The leather industry is looking forward to know whether sufficient raw materials are available in the near future domestically their business strategies. To get estimate for next few years to understand the trends in availability of hides and skins there does not exist any standard methods or software. The availability of hides and skins are highly dependent on the livestock population, cattle, buffalo, goat and sheep. Hence in order to estimate the availability of hides and skins, the livestock population for each of the category should be known. This paved the way to identify some statistical models to predict the livestock population based on the census data. Unlike the human population prediction, livestock projection requires additional attention, as slaughtering is a critical input for forecasting.

The first Indian livestock census was carried out in 1951. It was supposed to be carried out once in every five years. Unfortunately, we are having the census data in an irregular (lag) period. Even though census data depends on time, the time series modelling cannot be applied as such, as the data is unevenly spaced. Unevenly sampled time series are common in many real life situations. The vast majority of methods however can only handle regular time series and cannot be easily extended to unevenly sampled data. It is a common practice to ignore the times and treat the data as if it were regular. This will clearly introduce a significant bias leading to incorrect predictions. Jone (1985) proposed a state space representation and showed that for Gaussian inputs and errors, the likelihood of data can be calculated exactly using Kalman-filters.

Erodgan (2005) developed two time series model, one by assuming stationarity and the other by relaxing the stationarity assumption by allowing more general dependence on the current time, time difference and the state of process at a given time. The most commonly used methods that accommodate the unequal intervals in the time series are commonly used to estimate trend. The oldest and predominant method is a log-linear regression of counts against time, where slope of the regression gives the population trend (Caughly, 1977). This model is referred to as Exponential Growth Observation Error (EGOE) model, because it assumes the variability in data arises purely as sampling or observation error with the population itself governed by deterministic exponential growth. The second method for estimating exponential trend parameters assumes that the population is censuses (no observation error) and that

variability in abundance is entirely due to growth rate fluctuations caused by environmental variability or process noise. This model is referred to as the exponential growth process noise (EGPN) model. Jean et.al. (2009) explain how to obtain maximum likelihood estimates for EGSS model. Processing and analyzing unevenly spaced time series data in its unaltered form avoids the biases and information loss caused by data transformation. Eckner (2012) deals with analyzing unevenly spaced time series data without transformation. Even though methods are available to tackle the unevenly spaced time series data, unfortunately none of the above methods can be applied for predicting the livestock population. This motivated to develop a new statistical model to predict the livestock population and there by estimate the availability of hides and skins.

The article is divided into six sections. The information regarding the livestock census data and different methods for dealing with uneven sample are dealt with introductory section. The statistical modelling is discussed in Section 2, Section 3 deals with the software developed along with working model. Validation of the statistical model based on the software output in comparison with census data and linear regression models are given in Section 4. Finally Section 5 ends with concluding remarks.

II. STATISTICAL MODELLING

The source for predicting the livestock population is from the livestock census data. The possibility of existing methods and models has been reviewed and found that the existing models are not suitable for predicting the livestock population. As far as livestock population prediction is concerned, a critical parameter like off-take rate needs to be considered. In the usual models the off-take rate cannot be incorporated.

A. The Model

The model is formulated based on three factors namely livestock population, annual growth rate and net- off take rate (NR). The annual growth rate (g_t) is the weighted average of the growth rate calculated for each time interval. The weighted average is been used as the time gap between the census period as it is not uniform in nature. The net-off take rate is taken from CLRI survey report, Naidu et. al. (2005) and is calculated as follows,

$$NR = (Fallen\ Animals + Slaughtered\ Animals)/(Total\ Livestock)$$

Let y_t be the livestock population by time t, h_t the hides or skins available at time t. The availability of hides or skins can be calculated using the formula

$$h_t = y_t \times NR. \tag{1}$$

Hence the livestock population for the forecasted year can be modelled as,

$$y_{t+1} = y_t + (y_t - h_t)\, g_t + \varepsilon_t \tag{2}$$

where $\varepsilon_t = (y_t - h_t) * S.E.(g_{t_i}); i = 1, 2, ..., t-1; S.E.(g_{t_i}) = \text{Standard Error of the growth rates}$.

Moreover, ε_t follows asymptotically normally distributed with mean μ and variance σ^2 $\left(\varepsilon_t \square N(\mu, \sigma^2)\right)$.

From equation (2) one can forecast the livestock population over a period of years. It is similar to stochastic models in the sense that the current livestock population depends only on the just preceding year. The equation (1) is used to estimate the availability of raw materials; this in turn forms the input for predicting the livestock population. This statistical model is useful for making prediction and estimation manually. As far as industry and researchers in leather science are concern, this statistical model may not be comfortable to work with. Hence software has been developed for forecasting the livestock population and estimation of hides and skins. The software development is discussed in the following section.

III. SOFTWARE DEVELOPMENT

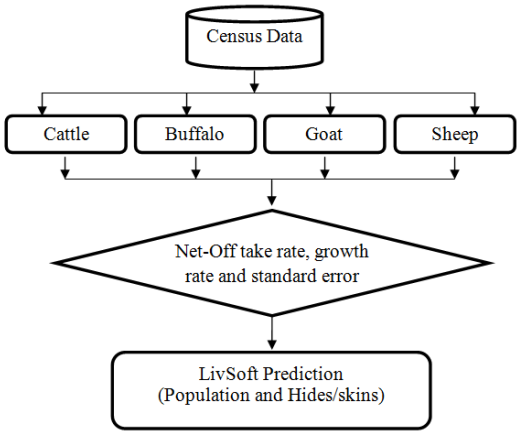
In recent past, most of the reports and bound volumes also called print resources were converted or developed as part of the information dissemination via World Wide Web (WWW) with value addition. In this contest, we report an interactive software model for forecasting livestock population and estimation of hides and skins. The LivSoft allow the users to choose various options to get the desired result with the user authentication.

This software has a data base on livestock population obtained from livestock census which was created using My SQL data base (Ben (2006)). The front end of the software was developed in JAVA (Budi Kurniawan (2002)). It automatically calculates the parameters like annual growth rate and standard error, which forms an input for the statistical model to predict the livestock population and thereby estimating the availability of hides and skins. This version will predict the livestock population at the national level.

The software is simple to use and flexible to choose the period from the data base to increase the accuracy of the output by incorporating the different trends. The user has to choose the category of animals from the “Category” box, the starting year from the drop down “Start Year” column as the base year, the end year in the “End Year” column and enter the year for which prediction is required in the “Predict Year” column, then click “Go” or click enter. As the output, the category of the animal, Net-off take rate, Annual growth rate and the prediction of livestock population in million and their corresponding hide or skin yield in million pieces will be displayed.

A. Working Model

The user will be getting two outputs; livestock population and availability of hides/ skins for cattle, buffalo, goat and sheep. The parameters like net-off take rate and growth rate are the two major factors influencing the prediction. At the same time the standard error of growth rate for that particular period were also computed.



B. The Livsoft Software

The home page of the LivSoft has a Login and About Us link. Login link will take us to login page, About Us will brief about the Department and its activity. The home page highlights about the software and its usage.

Like any other software the users will be authenticated with a username and password.

Figure 1: Software Home Page

Once click on the Login link, the login page will appear. In login page, one has to enter login name and password and click login icon to go to access page.

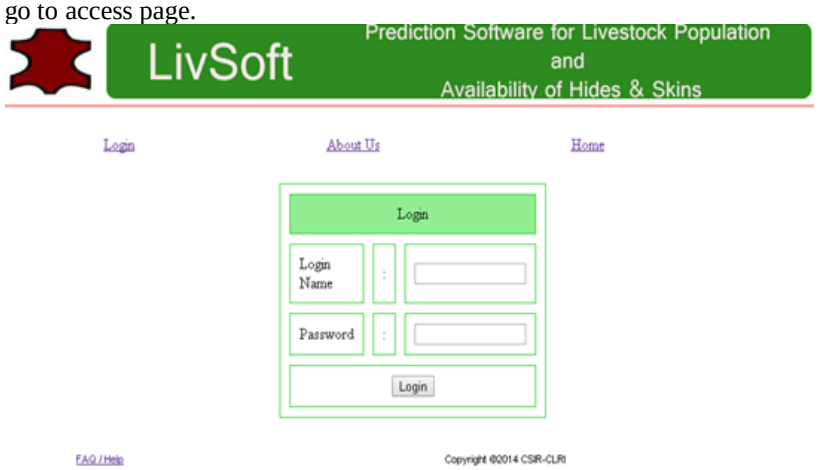


Figure 2: Login Page

Figure 3: Access Page

After clicking the access, it will take us to input page, where we have to choose the category, starting year, end year and prediction year. Then click Go, the predicted livestock population and availability of hides and skins for nation wise will be displayed in the output page (see Fig. 5).

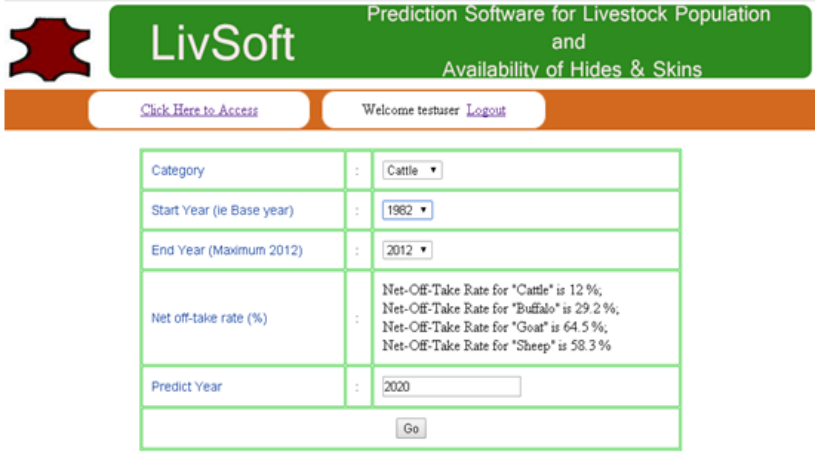


Figure 4: Input Page

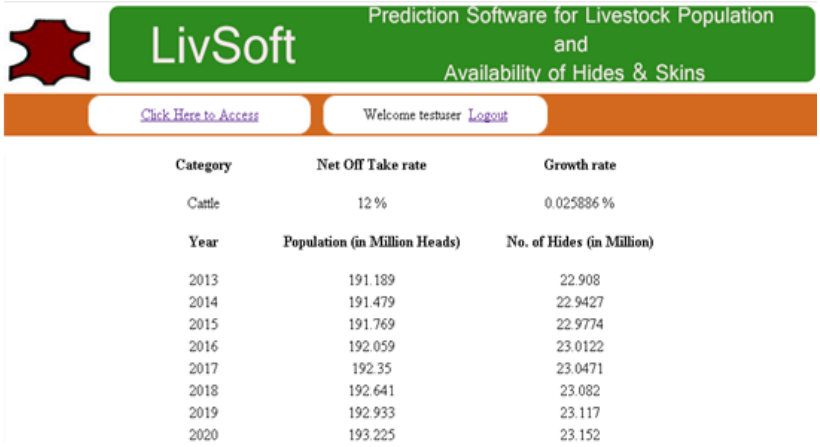


Figure 5: Out Put Page

C. Tools and Technologies Used To Develop the Software

The LivSoft was developed in J2EE with MySQL as backend in the tool kit Net Beans IDE 7.0.1. Java applications are typically compiled to byte code that can run on any Java virtual Machine (JVM) regardless of computer architecture. MySQL is a Relational Database Management System (RDBMS). The model algorithm for the software login page for cattle and variance calculation are given in appendix.

IV. COMPARISON AND VALIDATION OF THE MODEL

To validate the statistical model and the software, we considered the livestock census data from 1951 to 2007 and forecast is made for the year 2012. The forecast obtained from the software is compared with the actual census data in million heads to validate the software model and to find the accuracy of the forecast. The forecast is also compared with the linear regression models. Table 1 gives the validity and accuracy of the statistical model and software in comparison with regression models.

The fitted regression equations are as follows

$$Population_{Cattle} = 0.76 \times Y - 1326.82$$

$$Population_{Buffalo} = 1.11 \times Y - 2132.17$$

$$Population_{Goat} = 1.646 \times Y - 3168.76$$
$$Population_{Sheep} = 0.312 \times Y - 571.57$$

Table 1: Validation and Comparison of LivSoft (Values in Million Heads)

Pop. in Million Heads	Census data	LivSoft Model		Regression Model	
	2012	Value	% Variation	Value	% Variation
Cattle	190.9	201.09	5.33	210.3	10.19
Buffalo	108.7	106	-2.45	101.2	-6.91
Goat	135.2	142.17	5.17	148.4	9.79
Sheep	64.94	65.32	0.58	74.28	14.38

The percentage variation given in Table 1 shows the accuracy of the prediction and it is clear that the error is less than 6% for the forecast obtained from LivSoft. While the forecast obtained from regression models have double the variation compared to that of LivSoft predictive model. This justifies the significance of new statistical model and the usefulness of the simple software tool. The software guarantees minimum 90% accuracy.

V. CONCLUSION

The LivSoft software is user friendly software. It predicts the livestock population over a period of years with limited inputs and also estimates the availability of the hides and skins. This prediction is very useful to the leather industry, researchers and the ministry. This information helps the leather industry to plan appropriate strategies for improving their business, researchers for identifying the research problems and the ministries to frame appropriate policies to support famers as well as industries linked with livestock. Even though the statistical model there by software has been developed based on Indian scenario, it can be applied to the world livestock as well, provided the information regarding the net off take rate is available.

ACKNOWLEDGEMENT

The authors are thankful to CSIR- Central Leather Research Institute (CSIR-CLRI), Chennai for the financial support under the 12th five year plan (STRAIT - CSC 0201) project to carry out the study. The authors also acknowledge the CSIR-CLRI Communication No. 1154. This work has been done while the second (corresponding) author was working with CSIR – CLRI as Scientist.

REFERENCES

[1] Ben Forta, *MySQL crash Course*. Sams Publishing. USA, 2006
[2] Budi Kurniawan, *Java for the web with Servlet. JSP and EJB*. New Rider Publishing, USA, 2002
[3] G. Caughley, *Analysis of vertebrate populations*.-Wiley, 1977.
[4] A. Eckner, *A frame work for analysis of unevenly-spaced time series data*. Working Paper, 2012.
[5] E. Erdogan., S. Ma, A. Beygelzimer. and I. Rish, *Statistical models for unequally spaced time series*. Technical Report. IBM T.J. Watson Center, 2005.
[6] Jean-Yves Humbert, L.Scott Mills, S. H. Jon and Brain Dennis, A better way to estimate population trends. *Oikos*. 2009, 118, 1940 – 1946.
[7] R. H. Jones, *Time series analysis with unequally spaced data*. In E.J. Hannan, P.R. Krishnaiah and M.M. Rao (Eds.), *Time series in the Time Domain, Handbook of Stattistics*, North Holland. Amsterdam, 1985, 5, 157-178.
[8] B. K. Naidu. et.al., *Report of all India survey on raw hides and skins*. Technical Report. Vol. 1. CSIR-Central Leather Research Institute. Chennai. India, 2005.
[9] World Wide Web: Proposal for a Hyper-text Project. W3.org (1990-11-12). Retrieved on 2013-07-17. (http://www.w3.org/Proposal.html).

APPENDIX

Algorithm for Logic page (A few line of code written in J2EE)

```
NoofObservation = Math.sqrt(dif);
StandardError = SD / NoofObservation;
HideweightData = data * NetOffTakeRate;
Var1 = data - HideweightData;
CurrentYearData = data + StandardError * Var1;
data = CurrentYearData;
```

