

Influence of Rainfall on the Seed Production of Indian Major Carps-a Case Study of T.B.Board Fish Farm, T.B.Dam, Hospet

V.S.Kulkarni, Anandtheerth Puranik

Associate Professor and HOD, Department of Mathematics, N.V. Degree College, Gulbarga, Karnataka, India

Assistant Director of fisheries, T.B Board Fish Dam, T.B Dam, Hospet, Karnataka, India

ABSTRACT: Farmers everywhere depend on access to good quality seed, which is fundamental to their carp production systems. Most seed production change according to the monsoon variations in strengthening production capacities, yet farmers in developing countries still face difficulties in accessing the quality seed of the varieties that they require. A Case Study Was Conducted to know the extent of relationship between the Pre-monsoon and Post monsoon rainfall on the Seed Production of Indian Major carps Labeo and exotic carps for a period of 19 Years Pre-monsoon had a Positive and Highly Significant Correlation on the Seed Production of catla and also to the total Seed Production (P0.01). The Number of rainy days had a significant Correlation with Seed Production of Catla Followed by Rohu and Mrigal ((P0.01).A negative relationship was observed in case of common carp seed production and pre-monsoon rainy days based on the current data available over the period of 19 years. In this paper, we find the forecasting statistics for the seed production over the three monsoon periods for couple of years by the analysis of the given data using the non-linear regression modeling.

KEYWORDS: Regression, RBF (radial bias kernel), forecasting, kernel, Laplacian, polynomial, ANOVA.

I. INTRODUCTION

Breeding in fishes is regulated by environmental factors that trigger internal physiological mechanisms. The final event of the breeding cycle, the release of eggs and milt resulting in spawning can be controlled by placing the fish in an appropriate environment and by changing its internal regulatory mechanism by injecting hormones or other inducing substances. The internal mechanisms that regulate spawning are similar in most fish (1). The external environmental factors that control reproduction vary considerably among species (2).The availability of seeds of Indian major carps (IMC) comprising catla (catla catl Rohu (Labeo rohita) and Mrigal (Cirrhinus mrigala) and common carp (Cyprinus carpio) and other exotic carps in fish farm at the right time plays an important role in the development of fisheries of the region in the country. The objectives of the fish farm have been to produce maximum number of good quality seeds of carps to cater to the needs of reservoirs and tank fisheries of the States. But there has been a lacuna in the production and supply of seeds at right time due to climatic and environmental factors operating in the regions.

Karnataka State is blessed with lot of inland water resource consisting of 2500 major tanks (approx, 1.20lakh ha.), 3000 minor tanks (approx.1500 ha.) and 20reservoirs (approx, 1.20 lakh ha) as per fisheries department statistics of 2008-09. The present fish seed production of the state is around 20 crores of fry per annum, which is meeting only one third of the total requirement. The TB Board Fish Farm is one of the second largest seed production farm in the state contributing around 10 to 14% of the total seed production (fry) of the state. The rain fall plays an important role in breeding response and seed production of IMC and other carps. The amount of rainfall and the frequency of rainfall (number of rainy days) have a role to play as natural stimuli in changing physicochemical Characteristics of not only pond water but also the climatic conditions which facilitate gonadal development and breeding process of IMC, Hence. Current study is aimed at knowing the influence of rain fall on seed production of IMC and common carp over a period of two decades from 1991-92 to 2009-10 years since the reports on such works are scanty.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

II. MATERIAL AND METHODS

During the period of study in TB Board Fish Farm, The induced breeding trials have been conducted in 4 to 6 number of cement breeding tanks of size 3.5*6.2+1.2 m which had proper flowing water conditions. A cooling tower erected on one of the breeding tank is used to reduce air and water temperatures by 2 to 3 C during hot climatic conditions in pre monsoon and monsoon period to facilitate breeding.

On an average 45 to 55 kg of total female brooders per day were injected with reproductive hormones and released in 3 to 4 breeding tanks. A ratio of 1:1 between female and males was 10 to 12 mg. per kg and 4 to 6 mg per kg of body weight of fish respectively. The dose of Ova prim, Ova tide and Ova –FH injected for females and males were 0.4 to 0.5 ml per kg and 0.1 to 0.2 ml per kg of body weight of fish respectively. Chinese and Glass jar hatcheries are used for hatching the eggs.

All the brood stock and nursery management practices of fish farm regarding manuring, feeding etc., is as per the standard norms set by the Department of Fisheries, Government of Karnataka. Feed consisting of rice bran, Ground nut oil cake, powdered mixture of broken rice, maize, horse gram, soybean and fish meal in different combinations at 2% of body weight of brood stock have been tried over the years. Combination of organic and inorganic manures is practiced to get desired plankton production. The data was statistically analyzed using correlation coefficient for three different periods of monsoon viz., pre monsoon (March to may). Monsoon (June to September) and post monsoon (October to December) on the Indian major carps. Later correlation coefficient was tested for its significance by using t-test.

III. EXISTING SYSTEM

The study of rainfall and monsoon season effect on seed production of Indian major carps and exotic carps has been conducted for the duration of 19 years. The conduct of study on the rainy days and monsoon season had shown that these factors are the good measures for developing better seed production of Indian major carp. Pre-monsoon had a better correlation on the seed production of Catla and also to the total seed production. The number of rainy days had a significant correlation with seed production of Catla followed by Rohu and Mrigal. And adverse relationship was observed in case of common carp seed production and pre-monsoon rainy days.

The important factors required for analysis of seed production of carp is number of rainy days i.e. the amount of rainfall & the frequency of rainfall and monsoon seasons (Pre-monsoon, Monsoon and Post-monsoon). The factors other than rainy days and monsoon season which have impact on seed production are - the amount of injection of reproductive hormones, fish meal and atmospheric conditions like air temperature, water temperature, and pH of water.

A large variety of methods exist to increase prediction speed. Two main classes of approaches can be identified: linear regression for models and approximating the decision function of a given model. Our proposed approach belongs to the latter class. Complementary to those approaches, it is worth mentioning that older research attempted to increase evaluation speed for the exponential function because it represented a significant bottleneck for neural networks.

Regression analysis can be performed on the provided statistical data to estimate the relationship among variables. Regression analysis has many techniques for modeling and analyzing several variables, which means the importance is given to the relationship between a dependent variable and one or more independent variables. By using regression analysis one can get vivid picture of how the dependent and independent variables are related. For example the value of dependent variable can be declared as either increasing or decreasing depending upon the value of independent variable, because the dependent variable will obtain a value only when we provide a value to independent variable. The estimation target is a function of the independent variables called the regression function. The linear regression model for any given population of a dependent and independent variable can be formulated as:

$$y = \beta_0 + \beta_1x + \epsilon \quad (1)$$

Where

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

y - dependent variable, x - independent variable, β_0 - the intercept of the line, β_1 - the slope of the line, ϵ - statistical noise or error

The linear regression model provides how well the data points lie on the line. The line should be drawn in such a way that almost all data points should lay on this line. By looking at the regression line and arrangement of data points on the line specifies how well the regression model fits into the dataset.

To obtain linear regression model we must calculate these expressions:

Mean - is calculated by dividing the sum of observed values by the number of observations, n. Although data points fall above, below, or on the mean, it can be considered a good estimate for predicting subsequent data points. The formula for finding the mean of x and y quantities are given below

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (3)$$

Standard Deviation - measures the amount of variation or dispersion from the mean. The lower the standard deviation, the better the data statistics.

$$S_{xx} = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \quad (4)$$

$$S_{yy} = \sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n} \quad (5)$$

$$S_{xy} = \sum_{i=1}^n x_i y_i - \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n} \quad (6)$$

Where

$(x_i, y_i) = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$ is the input data . To calculate regression equation,
 $y = \beta_0 + \beta_1 x + \epsilon$

We have slope estimate as, $\beta_1 = \frac{S_{xy}}{S_{xx}}$ and intercept estimate as $\beta_0 = \bar{y} - \beta_1 \bar{x}$.

The slope of a line is a number that describes both the direction and the steepness of the line. If we find that the slope of the regression line is significantly different from zero, we will conclude that there is a significant relationship between the independent and dependent variables.

The y-intercept is the place where the regression line $y = \beta_0 + \beta_1 x + \epsilon$ crosses the y-axis (where $x = 0$), and is denoted by β_0 . The slope of a regression line is used with a t-statistic to test the significance of a linear relationship between x and y.

Correlation coefficient is a measure of the linear correlation i.e. dependence between two variables X and Y, giving a value between +1 and -1 inclusive. The value of correlation coefficient can be interpreted as - 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation. The correlation coefficient between data columns x and y is calculated using:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad (7)$$

From the above equation, the coefficient of determination R^2 can be computed by squaring correlation coefficient, the formula is given by:

$$R^2 = \frac{(S_{xy})^2}{S_{xx}S_{yy}} \quad (8)$$

This represents the proportion of the total sample variability in y that is explained by a linear relationship between x and y. R^2 value signifies how well the model fits into regression analysis. If the value of R^2 is close to 1 then the model fits well for regression analysis and if the value of R^2 is close to 0 then the model fits badly for regression analysis [3][4].

In linear regression analysis, we can obtain relationship only between a dependent and an independent variable, but in non-linear regression analysis we can obtain relationship between a dependent variable and one or more independent variables.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

IV. PROPOSED SYSTEM

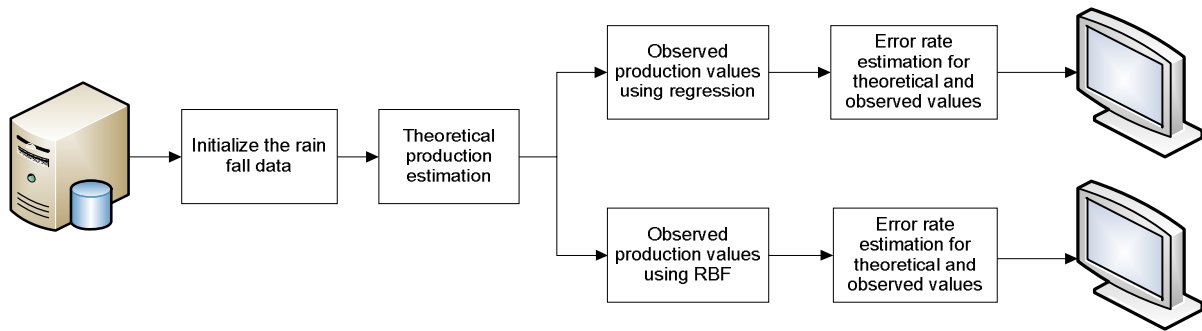


Figure 1: Block diagram of the proposed system

In statistics, regression analysis involves in estimating the association among variables which includes a dependent and one or more independent variable. To be more precise it helps in predicting how an independent variable varies the dependent keeping the other independent variables constant. In all cases, the estimation target is a function of the independent variables called the regression function. Regression function takes its part in both the classes of the regression analysis which include linear and non-regression function. Both the forms of regression however deal with the goodness of the fit. Essentially this means estimating how well the curve (i.e. the function) describes the data. The most commonly used measure of the goodness of fit is least squares. The least square method fits a flock of data points using a straight plane or flat plane. In most of the cases the data would be randomly plotted away (for example, if something is growing exponentially, which means growing at a steady rate). In such locations a straight plane cannot fit the data space thus reducing the prediction or classification accuracy. To fit something like this, you need non-linear regression. Often, you can adapt linear least squares to do this. The method is to create new variables from your data. The new variables are nonlinear functions of the variables in your data. If you construct your new variables properly, the curved function of your original variables can be expressed as a linear function of your new variables. Expressing non-linear data again into a linear data is probably not a good procedure to serve the need because of the lengthy simple calculations. However there are general methods which significantly accomplish the task of non-linearity. The concept of non-linearity can be accomplished in different ways based on the application required. For example if the data points are classified in a constantly changing rate, exponential equations can be applied. Certain type's nonlinear regression functions have served and will continue to serve as useful models for describing various physical and biological systems. We list a few of these situations for the case of a single predictor variable [5].

The following functions have considered in modeling the relationship between seed yield Y and the spacing between rows of plants, concentration Y of a drug in the bloodstream and time X after the drug is injected when this concentration is measure, the rate Y of a chemical reaction and the amount X of catalyst used, and many other relationships.

$$\mu_Y(x) = \frac{1}{(\beta_1 + \beta_2 x)\beta_3} \quad (9)$$

$$\mu_Y(x) = \frac{1}{\beta_1 + \beta_2 x + \beta_3 x^2} \quad (10)$$

$$\mu_Y(x) = \frac{1}{\beta_1 + \beta_2 x \beta_3} \quad (11)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

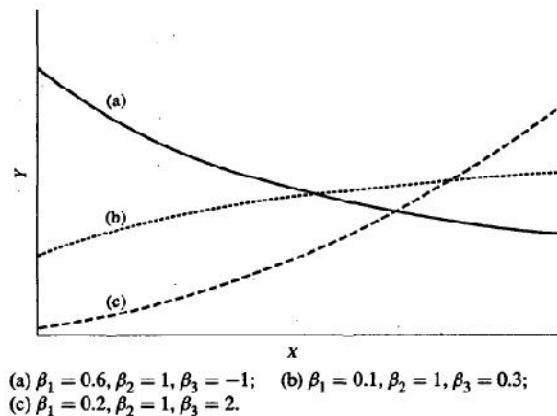


Figure 2: Three Members of the Family Curves

The family of curves obtained from figure. 2 by embedding the chain of values into the above equation leads to the curves which may look as in the above figure. The figure resembles the exponential family of curves. Since the regression data of our system systematically doesn't classify our needs properly. We look on to sinusoidal way of modeling the data. To do this a different kernel methods have to be introduced [6].

A. Kernel Methods

Kernel methods owe their name to the use of kernel functions, which enable them to operate in a high-dimensional, implicit feature space without ever computing the coordinates of the data in that space, but rather by simply computing the inner products between the images of all pairs of data in the feature space. This operation is often computationally cheaper than the explicit computation of the coordinates. This approach is called the "kernel trick". Kernel functions have been introduced for sequence data, graphs, text, images, as well as vectors. To be more specific the kernel trick can be specified to any algorithm which solely depends on the dot product between two vector points. Wherever a dot product is used, it is replaced by a kernel function.

When properly applied, those candidate linear algorithms are transformed into a non-linear algorithms (sometimes with little effort or reformulation). Those non-linear algorithms are equivalent to their linear originals operating in the range space of a feature space ϕ . However, because kernels are used, the ϕ function does not need to be ever explicitly computed. This is highly desirable, as we noted previously, because this higher-dimensional feature space could even be infinite-dimensional and thus infeasible to compute. There are also no constraints on the nature of the input vectors. Dot products could be defined between any kind of structure, such as trees or strings.

1. Polynomial Kernel

The Polynomial kernel is a non-stationary kernel. Polynomial kernels are well suited for problems where all the training data is normalized.

$$k(x, y) = (\alpha x^T y + c)^d \quad (12)$$

Adjustable parameters are the slope alpha, the constant term c and the polynomial degree d.

2. Laplacian Kernel

The Laplace Kernel is completely equivalent to the exponential kernel, except for being less sensitive for changes in the sigma parameter. Being equivalent, it is also a radial basis function kernel.

$$k(x, y) = \exp\left(-\frac{\|x-y\|}{\sigma}\right) \quad (13)$$

It is important to note that the observations made about the sigma parameter for the Gaussian kernel also apply to the Exponential and Laplacian kernels.

3. ANOVA Kernel

The ANOVA kernel is also a radial basis function kernel, just as the Gaussian and Laplacian kernels. It is said to perform well in multidimensional regression problems [7].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

$$k(x, y) = \sum_{k=1}^d \exp(-\sigma(x^k - y^k)^2) \tag{14}$$

B. Radial Basis Function For Non-Linear Regression Analysis

In regression analysis, the (Gaussian) radial basis function kernel, or RBF kernel, is a popular kernel function used in various learning algorithms. In particular, it is commonly used in vector classification. The widely used Radial Basis Function (RBF) kernel is known to perform well on a large variety of problems. RBF network can be used find a set weights for a curve fitting problem. The weights are in higher dimensional space than the original data.

Alternatively, it could also be implemented using the adjustable parameter sigma plays a major role in the performance of the kernel, and should be carefully tuned to the problem at hand. If overestimated, the exponential will behave almost linearly and the higher-dimensional projection will start to lose its non-linear power. In the other hand, if underestimated, the function will lack regularization and the decision boundary will be highly sensitive to noise in training data.

There are three basic classes of radial basis functions.

$$\phi(r) = (r^2 + c^2)^{1/2}, \text{ for some } c > 0 \text{ and } r \in \mathbb{R}$$

$$\phi(r) = \frac{1}{\phi(r) = (r^2 + c^2)^{1/2}}, \text{ for some } c > 0 \text{ and } r \in \mathbb{R}$$

$$k(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{15}$$

$$k(x, y) = \exp(-\gamma\|x - y\|^2) \tag{16}$$

Gaussian functions are probably the most used. In general, the selection depends on the application.

The above analogies prescribe to describe the data of a kernel trick used as a part of the regression and non-linear regression analysis. In the context of this our system would predominantly exploit the kernel on the regression systems to quantify the prediction analysis and their error rate. The error rate is a measure of the accuracy of predictions. Recall that the regression line is the line that minimizes the sum of squared deviations of prediction. In our context we embody the non linear regression line patterns in order to differentiate the carps based on the monsoon for the prediction.

V. RESULTS AND DISCUSSIONS

Due to the uncertainty in forecast applications, there is no unified standard for the solution. The cost of theoretical methods and tests is very high and the effectiveness of these tests has not been confirmed. To aim directly at uncertainty information and artifacts in information, the limitation of theoretical modeling can be overcome by the use of machine learning methods. Reasonable judgment, satisfactory predictions and ideal forecasts can be achieved by learning methods based on existing knowledge and experiences in solving problems.

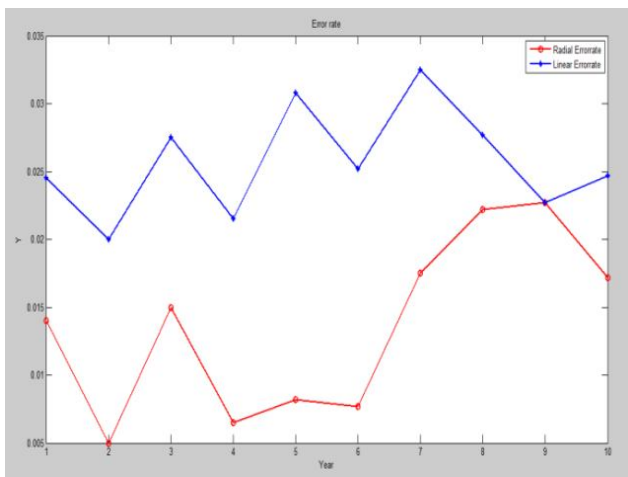


Figure 3: Prediction value curves

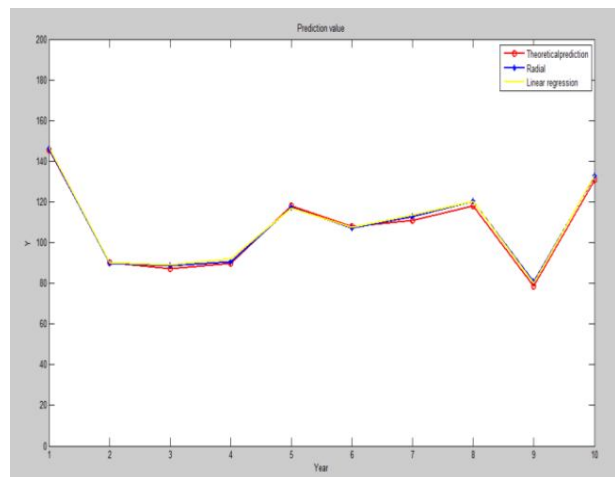


Figure 4: Error Rate Curves

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

The continuous growth and demand of food commodities and uncertain natural rainfall and the threat of population explosion strongly motivates to estimate the production analysis and consumption. Hence we produce a prediction table and their simulation analysis to verify the error rate and the prediction analysis of both RBF and regression kernels. The behavior of the table can be viewed in the table 1 and 2. The theoretically estimated results are compared with the regression analysis and RBF regression to predict the crap production rate as shown in figure 3.

The figure 4 estimates the error rate identified during the forecasting. The error rate is calculated taking the theoretical estimation as reference and subtracting the theoretical predicted values with the predicted values with the predicted values of the RBF and linear regression

Year	Rainfall in mm	Theoretical prediction				Observed prediction (Radial basis function)				Error rate
		Catla	Rohu	Mrigal	Common carp	Catla	Rohu	Mrigal	Common carp	
2000-01	602	161.4	269.3	75.5	76.5	162.2	271.6	74.1	75.6	1.4
2001-02	698	132	134	27	68	131	132	29	67	0.5
2002-03	415	171	96	54	28	173	97	55	30	1.5
2003-04	369	161.4	100.5	40.8	57	163.4	101.8	41.1	56	0.65
2004-05	664.6	215.4	162.4	63.5	31	216.4	161.5	61.1	30	0.82
2005-06	859.2	142.9	138.9	120	30	139.4	139.3	121	29	0.77
2006-07	365	123	221.1	80.52	19	126	223.4	81.22	20	1.75
2007-08	600.3	104.2	182.2	149.6	37	106.4	184.5	152.2	39	2.22
2008-09	656.7	116.45	102	89	6	118.6	104	92	9	2.27
2009-10	742	176	256.2	91	1	178	257.1	93	3	1.72

Table 1: Prediction Table Using Linear Regression and RBF

Year	Rainfall mm	Theoretical prediction				Observed prediction (Linear regression)				Error rate
		Catla	Rohu	Mrigal	Common carp	Catla	Rohu	Mrigal	Common carp	
2000-01	602	159.4	268.3	74.5	75.5	162.2	271.6	76.1	77.6	2.45
2001-02	698	130	131	26	66	132	133	27	69	2
2002-03	415	170	95	53	28	174	98	55	30	2.75
2003-04	369	161.4	100.5	40.8	57	163.4	103.8	42.1	59	2.15

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

2004-05	664.6	217.4	164.4	65.5	33	215.4	160.5	62.1	30	3.08
2005-06	859.2	143.9	140.9	123	32	140.4	139.3	121	29	2.52
2006-07	365	123	220.1	80.52	18	126	224.4	83.22	21	3.25
2007-08	600.3	103.2	181.2	150.6	35	105.4	184.5	153.2	38	2.77
2008-09	656.7	115.45	101	89	5	117.6	103	92	8	2.27
2009-10	742	175	255.2	90	2	178	257.1	93	4	2.47

Table 2: Error Rate Analysis Table Using Linear Regression and RBF

VI. CONCLUSION

The amount and frequency of rainfall of the monsoon periods have a definite role to play in the success of the seed production of IMC. However, the present circumstances of the unusual rainfall will drive us to depend on the concept of forecasting. Since forecasting does not provide accurate results, error rate or slope correction becomes mandatory. In this paper, we have introduced a method to construct an approximate forecasting using the non-linear regression (RBF kernel) and linear regression and estimated that RBF performance is far superior in prediction accuracy than linear regression.

ACKNOWLEDGEMENT

We would like to acknowledge the support of UGC (MRP (S_0424/13-14/KAGU005/UGC-SWRO)) in providing us the grant for the research.

REFERENCES

1. "Rottmann ,shireman, chapman N :1991", Hormonal control of reproduction in fish for induced spawning SRAC Publication No 424.
2. "Devendra mohan and Devendra Choudary: 2010", Variations in physic-chemical and microbiological characteristics of water during breeding of cyprinus carpio in a closed hatchery system. Journal of environmental biology, 31, 301-306.
3. "Gary Simon", SIMPLE LINEAR REGRESSION "New York University, Stern School of Business", Documents prepared for use B01.1305, 2004.
4. Abdullah A. Smadi and Nour H. Abu-Afouna", On Least Squares Estimation in a Simple Linear Regression Model with Periodically Correlated Errors, Austrian Journal of Statistics, Vol. 41 (2012), No. 3, 211-226
5. Angus M. Brown,"A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet" Department of Neurology, Box 356465, University of Washington School of Medicine, Seattle, WA 98195 -6465, USA
6. GK smyth,"Non-linear Regression".
7. CR Souza, "Kernel Functions for Machine Learning Applications", 2010.

BIOGRAPHY

Venkatesh kulkarni, has received M.Sc. degree in Mathematics from Gulbarga University, Karnataka, India in 1987. He has vast experience of 25 yrs in teaching U.G students. He Exiled Ph.D. Course work with Distinction (2011) & pursuing Ph.D which is expected to submit the Thesis in 2015. He was B.O.S. and B.O.A.E member from 2011-2014 and Organized Two Days Workshop from 28 and 29th January 2014 in the department of Mathematics Gulbarga University, Kalaburagi (Gulbarga) to revise the syllabus of B.Sc.course for U.G. lectures.

Anandteerth Puranik, is presently working as a curator in T.B. Board Fish Farm, T.B. Dam, Hospet (Karnataka) India. He is specialized in Master of fishing science (M.F.Sc.)