

# Predicting the Land Value using Regression Techniques and Artificial Neural Network

Velumani P, Nampoothiri N, Kavithra P

**Abstract:** Land value can be an important factor which influences the cost of construction on working in the project. The land has socio-economic and environmental values and the confronted problems on land involves the increasing costs for developing the land such as built up, agricultural, residential, commercial and industrial areas. Hence this paper concentrates on prediction of land value by considering some important factors that affects it. The study area has been selected under Tirupur district, being a developing one in Tamil Nadu. The eleven areas in four different taluks under Tirupur district were chosen for research work. The average values of monthly variation are taken for the chosen factor for the years from 2001 to 2017. Using regression analysis and artificial neural network, the prediction has been done for the future land value. The performance of both the model executed good and fit for forecasting results. Though both the model showed better results, Artificial Neural Network (ANN) showed accuracy than regression method.

**Keywords:** Land value, Regression model, Artificial Neural Network, Google images, Historical data

## I. INTRODUCTION

Urban planning is an important factor for developers to achieve sustainable development. Planning considers construction cost as a major problem due to the trending land price (Wang, 2017). Hence land value has to be studied for future decision making on planning a construction project. To estimate the land price and its trend in future, factors affecting the land price have to be studied (Sampathkumar, Forecasting the Land Price using Statistical and Neural Network Software, 2015) and for that real estate system can deliver the required data but it is a fluctuating random process. Hence guideline value for the selected area is collected over previous years. Land value can be dependent on various factors based on its location importance (Mukhlshin, 2017). In general suburban centers are getting developed which can be achieved by the process of suburbanization and grouping of retailers and offices that makes the land value to be increased at center of the city

(Ping, 2005) and also housing is an important factor which is influenced by population growth while this results on land prices to be increased based on the different use of land. The land use regulations might have a direct impact on property values based on the land use change pattern and also the governmental policies and demographic changes over the period of time. (Jaeger, 2006). The various factor that causes the land price has been collected through literature survey. The factors such as GDP, CPI, construction cost index, population, inflation rate is considered for the prediction of land market value (Wang, 2017). In general, the interest rates and inflation affect the land price which is influenced by cash rent payments. (Harmon, 2004). GDP and lending rates maintain a positive relationship with housing price based on the enactment of government controls (Guo, 2013). Also, another 12 factors were identified which affects the residential land value on prediction it for the district Klongluang in Thailand and finalized that only six factors showed great impact on it and listed them as area, location, zoning, surface, width and user (Khumpaisal, 2017). For testing and training models, time series data were used. Meanwhile the study has been carried out in Malaysia for also compared regression technique (Radzi, 2012). Gamma test, Neural network, time series forecasting was used for predicting residential property price in UK (Wilson, 2002). A comparison has been carried out between MRA and ANN on predicting housing values in the Rutherford County, Tennessee (Nguyen, 2001) and based on the training data sample and size, results are given which is better. For testing and training ANN model, the software used is Stuttgart Neural Network Simulator (SNSS 4.2). For evaluating and interpreting the model of multiple regression, the database has been introduced in the form of variables in SPSS software (Ping, 2005) and meanwhile the constant and the coefficient has been achieved in the process. A prognostic model has been created with the help of ANN for the estimation of real estate prices in EU market and it resulted a satisfactory degree of accuracy considering the various factors for the analysis (Cetkovic, 2018).

## II. METHODOLOGY

The procedure for the prediction of land value can be followed using the step by step analysis given below,

- To know the impact of land value and study what are the factors that control it

**Revised Manuscript Received on December 5, 2019**

\* Correspondence Author

**Velumani P\***, Civil department, Kalasalingam Academy of Research and Education, Krishnankoil, India Email: velumani7@gmail.com

**Nampoothiri N**, Civil department, Kalasalingam Academy of Research and Education, Krishnankoil, India Email: drnvn@gmail.com

**Kavithra P**, Civil department, Kalasalingam Academy of Research and Education, Krishnankoil, India. Email: ajkavithra@gmail.com

## Predicting the Land Value using Regression Techniques and Artificial Neural Network

- To collect the guideline values of the study area for the previous years and knowing its fluctuations
- To collect all the data of influencing factors and data analysis should be carried out
- To identify the most highly correlating factor (more than 0.75)
- To input the correlating data into the prediction tool
- To use the data for multiple regression analysis
- To use the data in artificial neural network model
- To compare the output of both the testing model
- To conclude which model delivers a better and accurate result

### III. STUDY AREA

Tirupur is district of Tamil Nadu which is an industrialized and developed area. The district is formed on 2009 that is separated from Coimbatore district and part of Erode district. The total extent of the district is about 516.12 sq.km. The district is bounded by Erode district on the North and northeast, Karur district in the east, Dindugal district in the south east and in the west by Coimbatore district. Climate is moderate because the district lies on the western part of Tamil Nadu bordering the western Ghats and the region enjoys maximum rainfall. The major rivers passing the district are Noyyal and Amravati. The Amravati river is the main source of irrigation over the region. Due to the availability of sources and comfort ability, the population keeps on increasing. The study area satisfies mostly in parts of industrial need and development in economy of the state. It comprises of seven taluks namely Avinashi, Madathukulam, Palladam, Udumalpet, Tirupur, Dharapuram and Kangeyam. In these some of the taluks are developing a lot after the announcement of separate district. Here out of seven taluks, four taluks were selected as study area for the prediction of land value namely Avinashi, Kangeyam, Palladam and Dharapuram.

The cities of Tirupur experiencing a leading growth rate in economics and it was achieved only when the government anticipates sufficient infrastructure and support to the clothing and knitting industry (times of India, 2018). Over the last decade, Tirupur has resulted strong economic growth. With a rapidly growing hosiery industry and an excellent export demand for quality garments, the town has been in the development of the textile trade. There are about 4720 industrial units in Tirupur town and many of them are knitwear and related industries (NITTTR, 2014).

#### A. Nearby places of the study area





##### 1) Avinashi

The nearest airport is the Coimbatore International Airport, which is about 32 km away from Avinashi. The nearest Railway Station is at Tirupur, which is about 15 km away from Avinashi

##### 2) Palladam

The nearest airport is Coimbatore International Airport about 35 km south of the city from Palladam via Sulur & L&T bypass Road. The nearest Railway Station is at Tirupur, about 16 km away from Palladam.

TABLE 1: DETAILS ABOUT THE STUDY AREA

Taluk	Google map	Co-ordinates	Area (sq. km)
Avinashi		11°11'43"N 77°16'7"E	11.65
Palladam		10.98°N 77.3°E	19.42
Dharapuram		10.73°N 77.52°E	7.02
Kangeyam		11.0054°N 77.5620°E	22.17

##### 3) Dharmapuram

The town is located at the distance of 28 km from the district headquarters Tiruppur, 45 km from the Erode and 70 km from the industrial town of Coimbatore. Sivanmalai is a Hindu temple located on the State Highway 172 and it is about 4 km from Kangeyam. Orathuppalayam Dam, built across the Noyyal river, is about 14 km from Kangeyam

##### 4) Kangeyam

The nearest railway station is Palani (35 km) and Tiruppur 48 km respectively. The nearest airport is Coimbatore International Airport (80 km).

#### IV. DATA COLLECTION

Land price influencing factors has been identified and shortlisted through the literature review and as follows are GDP (Gross Domestic Price), inflation, dollar rate, gold rate, silver rate, crude oil price, population, time and street count. The data for guideline values were collected between the years 2001 to 2017 which were the government records and through state government website (tnreginet). There were totally 11 major areas in the four taluks namely Avinashi, Kangeyam, Palladam and Dharapuram, were chosen under study area. For these major areas in taluk, the streetwise guideline values have been collected. The overall average of each area is taken into account. The data for other factors were collected through the indiastat website which provides the year by year data in the monthly basis(indiastat). The average value of monthly data for each year is taken and used for the analysis. The time factor was considered in order of years from 2001 to 2017. The area of development under the selected study area has been calculated based on the increased streets in each area and it is chosen as a factor for predicting land value and it is collected through official website of Tamil Nadu government.

GDP has been collected as year wise data for all the 11-study area. Inflation is collected as monthly data for each year and for each year average value of month wise data is calculated for the study area. Out of 216 data for inflation 18 data is consolidated for the analysis. Dollar value is considered from daily change rate and mean value for each year is computed to be used as a factor. Likewise, dollar value is collected, changes in gold rate & silver rate has been analyzed and year wise data is generated and also for the crude oil data. Street wise data which is chosen for the analysis has been available for 5 consecutive years from the governmental record. Also, as per the census report, only the data in the year 2001 & 2011 is available. The missing data for the in between years has been identified with the help of time series analysis and some forecasting formulas has been adopted for every 11 areas under study.

As population increases land value increases proportionally. The land values are not directly related with population but because of agglomeration economies. Agglomeration is the result of population change and increase of population causes people to chase the same land thus land value increases well. Thus, the result of chasing gradually enhances the rate of change in land value. (Bettencourt, 2013). Crude oil has an indirect impact on land value. The consecutive change of crude oil rate has its own effects on logistics industry which also inflation and transportation cost. If crude oil rate increases, some companies will manufacture products in such a way to minimize the transportation. Hence transportation automatically get reduced that affects the land use development (QuickFreight Quote, 2019). Land use has is affected due to the changes in transportation cost. Since accessibility is the main criteria to improve the land based on its location. Transport facility improves the economic growth of the society which is

connected to development of land use factors. As a result, urban expansion prevails that improves the population and further the land value is affected. (Litman, 2019). Sampathkumar (2015) has chosen gold price and silver price as a factor for prediction of land price and finally the above-mentioned factors are selected for data analysis.

#### V. DATA ANALYSIS

The collected data correlation is carried out for each factor with GLV. Some of the missed data were determined through forecasting formula and data analysis were conducted for attaining the correlation accuracy. The factors which are correlating the most were found out through repeating analysis of data. The highly correlated factors were selected and used for testing and training models in multiple regression analysis and artificial neural network. For getting better results, analysis has been carried out to find the most correlated year that may be the major cause since dependent variable can be highly correlated with previous year's value (Velumani, 2018). The general formula for correlation is given as,

##### A. Regression Analysis

The often method used for prediction is found to be regression analysis. Regression analysis is utilized to comprehend which among the independent variables are related to the dependent variables, and to establish the relationship forms (Sampathkumar, Forecasting the Land Price using Statistical and Neural Network Software, 2015). To enhance, more autonomous factors are included which can decide an assessing condition that portrays the relationship with more accuracy. Multiple regression analysis is a worthwhile technique used for analyzing the unknown value of a variable from the known value of two or more factor which are also called the predictors. Multiple regression is an approach for modelling the relationship between dependent variable and a greater number of independent variables. There may be only one independent variable in case of linear regression while MRA deals with number of variables which is the most common applications in defining a variable characteristic (Khamis, 2014).

The general regression equation is  $Y = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$

Where Y – Guideline value (INR) – Dependent variable

$b_1, b_2, b_3, b_4, b_5$  are the coefficient of the independent variables

a – Error constant  $x_1$  – Time (year)

$X_2$  – Population (Number)  $X_3$  – Gold rate (INR – 10gm)  $X_4$  – Silver rate (INR – 1 Kg)

$X_5$  – Crude oil (INR – 1 Barrel)  $X_4$  – Inflation (INR - Percentage)  $X_5$  – Street count (Number)

Here it was considered that guideline value as dependent variable and the other factors were as independent



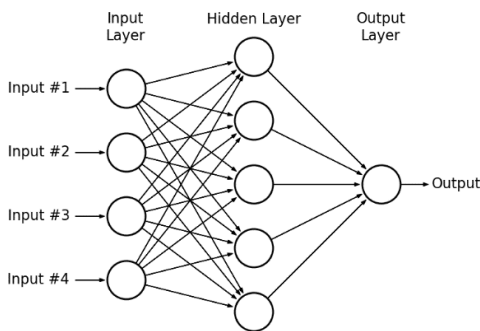


variables. The multiple regression formula was attained through SPSS software has been extracted.

**B. Artificial Neural Network**

To make the analysis in a simple way and quicker, neural network serves the need. It is capable of handling nonlinear problems (Ge, 2004). ANN is a valuation technology which surpasses the neural network of human brain comprises of number of processing elements that can be stored in its local memory (Sampathkumar, Artificial Neural Network Modeling of Land Price at Sowcarpet in Chennai City, 2010). The group of simple units, called as nodes, relate to each other by the edges, calls as synapses, which together to form neural network that can be organized in layers (Mimis, 2013).

NN comprises of processing nodes and connections among them. Each node in the layer connected with each node in the next layer by the means of connection. The structure of neural network is formed with an input layer, number of hidden layer and the output layer which process lot of data analysis and results accurate value for prediction (Khamis, 2014).



**Fig 1: Function of Artificial neural network (source: Artificial neural network – Textbook)**

**VI. RESULT AND DISCUSSION**

It is a difficult task to select validating data for carrying out forecasting. More number of researchers opined about fixing the validation and Calibration data. Akindele (1990) suggested that 80% of data has been taken for model calibration and 20% of data has been taken for model validating. 20% of whole data (18) is  $0.2 \times 18 = 3.6$  Number, approximately 3 data are used for validation and balance 15 data are used for calibration. Velumani 2018 opined that inverse the square root of whole data percentage value use to select the validation data ( $1/\sqrt{N} - 1/\sqrt{18} = 23.5\%$ ;  $0.235 \times 18 = 4.23$  Number).

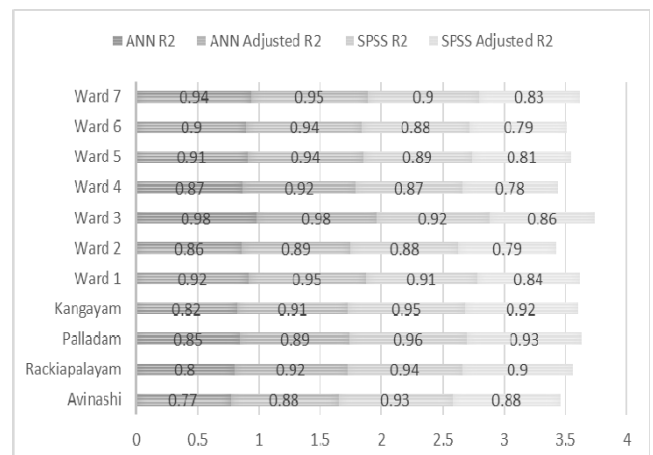
If validating sample get reduced, the result will be error free as required and this basic concept is also adopted for Artificial Neural Network for validation, thus it is helpful to attain minimized error. Co-efficient of determination ( $R^2$ ) value must be between 0 and 1, For industry-oriented prediction purpose

0.75 is the minimum value. This selected calibration and validation model is apt to predict the future, because  $R^2$  value

is more than 0.75 (Ref below table 2). More accuracy prediction happens when Adjusted  $R^2$  must be equal or higher than to the  $R^2$ . Here Regression analysis is strong the criteria of  $R^2$  and and weak in Adjusted  $R^2$  as compared to the Artificial Neural Network values.

**Table 2: Ann Co-Efficient of Determination Compared with Regression**

Study area	ANN		Regression analysis	
	$R^2$	Adjusted $R^2$	$R^2$	Adjusted $R^2$
Avinashi	0.77	0.88	0.93	0.88
Rackiapalayam	0.8	0.92	0.94	0.9
Palladam	0.85	0.89	0.96	0.93
Kangayam	0.82	0.91	0.95	0.92
Ward 1	0.92	0.95	0.91	0.84
Ward 2	0.86	0.89	0.88	0.79
Ward 3	0.98	0.98	0.92	0.86
Ward 4	0.87	0.92	0.87	0.78
Ward 5	0.91	0.94	0.89	0.81
Ward 6	0.9	0.94	0.88	0.79
Ward 7	0.94	0.95	0.9	0.83



**Fig 3: ANN Co-efficient of determination compared with Regression analysis**

**Table 3: Regression Formula in area wise**

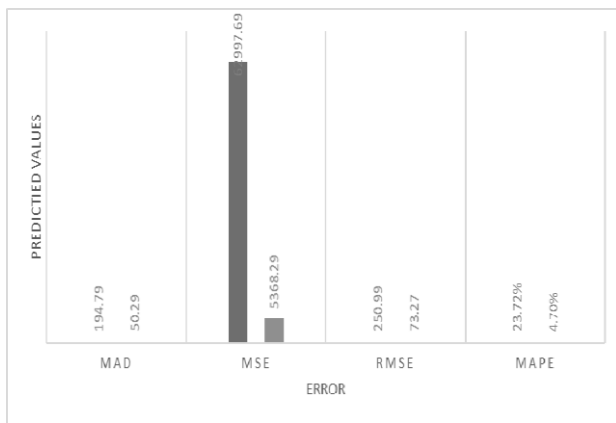
Study area	Regression formula in area wise
Avinashi	$Y = -9.52X_1 + 0.012X_3 + 0.003X_4 + 1.095X_5 - 6.901X_6 + 19826.59$
Rackiapalayam	$Y = -0.61X_1 + 0.004X_3 + 1.24X_5 + 2.47X_6 + 1132.03$
Palladam	$Y = -25.94X_1 + 0.004X_3 - 0.003X_4 + 2.074X_5 + 17.645X_6 + 49924.61$
Kankeyam	$Y = -37.99X_1 + 0.006X_3 - 0.002X_4 - 0.38X_5 + 6.78X_6 + 75600.41$
Ward 1	$Y = -1.03X_1 + 0.01X_3 + 0.003X_5 - 39.07X_6 + 2487.17$
Ward 2	$Y = -12.17X_1 + 0.03X_3 + 0.001X_4 + 0.26X_5 - 42.57X_6 + 24884.98$
Ward 3	$Y = 4.78X_1 + 0.04X_3 + 0.002X_4 - 0.24X_5 - 66.84X_6 - 7739.34$

Ward 4	$Y = -19.74X_1 + 0.05X_3 + 0.001X_4 + 0.41X_5 - 20.90X_6 + 40306.88$
Ward 5	$Y = -15.62X_1 + 0.06X_3 + 0.002X_4 + 0.27X_5 - 67.75X_6 + 33636.02$
Ward 6	$Y = -17.38X_1 + 0.06X_3 + 0.002X_4 + 0.32X_5 - 12.72X_6 + 35218.34$
Ward 7	$Y = -9.55X_1 + 0.06X_3 + 0.003X_4 + 0.09X_5 - 6.44X_6 + 19183.49$

Selected calibration data was calibrated and framed the equations (Ref: table 3) to predict the future value of guideline value in area wise.

**Table 4: Error comparison between Regression analysis and Artificial neural network**

Error	Regression analysis	Artificial neural network
MAD	194.79	50.29
MSE	62997.69	5368.29
RMSE	250.99	73.27
MAPE	23.72%	4.70%



**Fig 4: Error comparison between Regression analysis and ANN**

**Table 5: Area Wise Error Comparison for Tirupur (Study Area)**

Study area	Types of Error			
	MAD	MSE	RMSE	MAPE
Avinashi	46.9	2501.2	50	11.90%
Rakiya	27.79	1156.1	34	16.60%
Palladam	50.4	5273.1	72.6	9.20%
Kangeyam	45.5	4258.1	65.3	15.90%
Ward 1	15	254.9	15.9	11.50%
Ward 2	41.7	3451.3	58.7	11.30%
Ward 3	31.7	1256.8	35.4	7.70%
Ward 4	75.3	10705.7	103.46	21.10%
Ward 5	72.1	10903.6	104.4	9.80%
Ward 6	78.8	11798.6	108.6	16.70%
Ward 7	67.9	7491.2	86.5	7.90%

## VII. CONCLUSION

Using the empirical error measuring criteria, the performance of prediction between MRA and ANN is compared. Data model shows accuracy and fluctuates based on variation in specification and training size ANN showed better results than MRA on comparing the selected training size. It is very important to choose the method to be adopted for forecasting based on data which is under validation and training. The accuracy of ANN on compared to MRA is due to the uniformity in the data chosen in training size. The collected data vary once in five years must use the same value repeatedly and hence uniformity in the training size exists. This has an impact on MRA which shows less accuracy. However, ANN performed a better forecasting even though the repeated variables are used on small training size. Overall MAPE value showed 4% in case of ANN whereas 23% in MRA which can be due to the existence of uniformity in the size of data sample under functional specification. To conclude that based on the prediction result discussed above, one can be able to analyses the future land value which serves a part to control the construction price for proper planning and execution. Hence the prediction through ANN showed a best result that can be adopted for land value prediction which has uniformity in data handling in training size.

## REFERENCES

- Bettencourt, L. M. The Origin of Scalings in Cities. Science, 2013, 1438-1441.
- Cetkovic, J. Assessment of the Real Estate Market Value in the European Market by Artificial Neural Networks Application. Complexity, 2018, 1-10.
- Ge, X. J. Modeling Property Prices Using Neural Network Model for Hong Kong. International Real Estate Review, 2004, 121-138.
- Guo, M. The Empirical Analysis of Affecting Factors of Shanghai Housing Prices. International Journal of Business and Social Science, 2013, 218-223.
- Harmon, J. (2004). Effects of Inflation and Interest Rates on Land Pricing.
- indiastat. (kein Datum). Von indiastat.com: <https://www.indiastat.com/> abgerufen
- Jaeger, W. K. (2006). The Effects of Land-Use Regulations on Property Values.
- Khamis, A. B. (2014). Comparative Study on Estimate House Price Using Statistical and Neural Network Model. International Journal of Scientific & Technology Research 2014, 126-131.
- Khumpaisal, S. (2017). A Prediction of Residential Land Value, Case Study of Klong Nueng Subdistrict, Klongluang District, Pathumthani. International Conference on Innovations in Engineering, Technology, Computers and Industrial Applications, (S. 21-25).
- Litman, T. (2019). Evaluating Transportation Land Use Impacts. Victoria Transport Policy Institute.
- Mimis, A. (2013). Property Valuation with Artificial Neural Network: the case of Athens. Journal of Property Research, 128-143.
- Mukhlisin, M. F. (2017). Predicting House Sale Price Using Fuzzy Logic, Artificial Neural Network and K-Nearest Neighbor. International Conference on Informatics and Computational Sciences, (S. 171-176).
- Nguyen, N. (2001). Predicting Housing Value: A C comparison of Multiple Regression Analysis and Artificial Neural Network. JRER, 313-336.
- NITTTR, C. (2014). Slum free city plan of action- tiruppur corporation. Chennai: NITTTR.
- Ping, A. (2005). Residential Land Value Modelling, case study of Hankou, China. The Netherlands: International Institute for Geo- Information Science and Earth Observation Enschede.
- Quick Freight Quote. (2019). Von [www.directdrivelogistics.com](http://www.directdrivelogistics.com) abgerufen
- Radzi, M. S. (2012). Forecasting house price index using artificial



- neural network. International Journal of Real Estate Studies, 43-48.
18. Sampath Kumar, V. (2010). Artificial Neural Network Modeling of Land Price at Sowcarpet in Chennai City. International Journal of Computer Science & Emerging Technologies, 2010, 44-49.
  19. Sampathkumar, V. Forecasting the Land Price using Statistical and Neural Network Software. Procedia Computer Science, 2015, 112-121.
  20. times of India. (8. December 2018). Von m.times of india.com: [https://www.google.com/amp/s/m.timesofindia.com/city/coimbatore/irupur-looks-to-govt-for-growth-boosters/amp\\_articles/66994637.cms](https://www.google.com/amp/s/m.timesofindia.com/city/coimbatore/irupur-looks-to-govt-for-growth-boosters/amp_articles/66994637.cms) abgerufen
  21. tnreginet. (kein Datum). Von [www.tnreginet.net/guideline\\_value.asp](http://www.tnreginet.net/guideline_value.asp) abgerufen
  22. Velumani, P. Predicting India's Cement Price: A Comparison of Regression Analysis, Trend Analysis, And Artificial Neural Network. International Journal of Civil Engineering and Technology, 2018, 1907- 1915.
  23. Wang, L. Prediction on Land Market Value Based on the Real Estate market in USA. Journal of Mathematics and Statistics, 2017, 143-151.
  24. Wilson, I. Residential property price time series forecasting with neural networks. Knowledge-Based Systems, 2002, 335-341.
  25. Wisemen, S. (1966). correlation method. Manchester: University of Manchester.
  26. G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (references)
  27. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
  28. I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
  29. K. Elissa, "Title of paper if known," unpublished.
  30. R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
  31. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
  32. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

### AUTHORS PROFILE



**Velumani P**, M. Tech (Construction Engineering and Management) Pursuing Ph. D in the fields of prediction profile, Published 6 Nos of international journals. Working as a Civil Engineering faculty since 2014 at Kalasalingam Academy of Research and Education, Krishnankoil, India.



**Nampoothiri N**, Ph. D, Published around 50+ no of International journals. Finished Ph. D at Indian institute of Technology, Bombay. Recipient of workshop grant from NRDMS, DST



**Kavithra.P** M. Tech (Construction Engineering and Management) Published 2 Nos of international journals Studied at Kalasalingam Academy of Research and Education, Krishnankoil, India.