International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 3 (2017) pp. 283-288

Using Multiple Regression Analysis Lineal to Predict Occupation Market Work in Occupational Hazard Prevention Services

A. Guardiola Mouhaffel[1*], C. Martínez Domínguez [1], B. Arcones[1], F. Morán Redonda[1] & R. Díaz Martín[2]

[1] Camilo José Cela University, Madrid, Spain.
[2] Distance university of Madrid (UDIMA), Madrid, Spain.

** General Manager Grupo Marco Senegal, PhD Professor (Cela Open institute)*
Aggregate center of Camilo Jose Cela University.

Abstract

A simple model for Employability has been described. The model is simple due to the fact that it uses simple mathematical equation using Multiple Linear Regression (MLR) equations that can be easily understood. A worker data at a particular station is recorded by pooling data process .This study is the result of a survey of 6000 workers in occupational hazard prevention services (OHPS). The pool is highly representative, given that it considers The five main professions dedicated to this labor sector over 2010 and 2015, obtaining a fixed photograph of the temporal reality cited, since the surveys were carried out in the Basque Country, Aragon, Valencia, Andalusia, Catalonia, Madrid, Galicia, the Canary Islands, Castilla - La Mancha and Castilla – León. Engineering, Architecture, Socials Sciences, Medicine and Humanist Science are scoping his employability, polling workers numbers by profession and analyzing a predicting model using simple mathematical equation using Multiple Linear Regression (MLR). The employment rate is 71.80 % and the health sector is the most dedicated to this sector.

The development phase of the model is to obtain MLR equations using input set and output parameter. The coefficients of these regression equations have been used to estimate the future weather conditions. The personal computer and simple data processing software like MS Excel can be used to make and validate the model by the user itself. The results obtained show that MLR model can estimate the weather conditions satisfactorily.

Keywords: Polling workers, Employability, Multiple Linear Regression, Occupational hazard Prevention Services

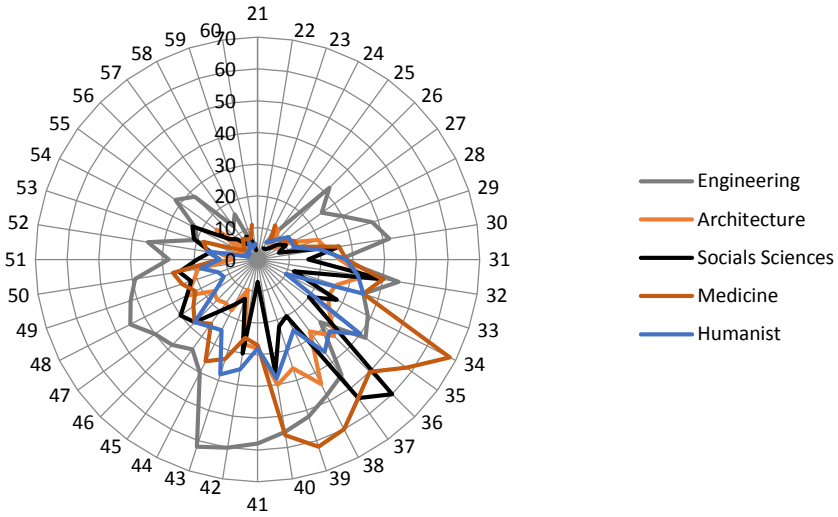

Fig 1:Age multiprocessing polling workers OLRP

INTRODUCTION

The aim of this paper is created a model to predict the labor market overall occupation labor in occupational hazard prevention services (OHPS) by occupational labor .International labor organization (ILO), the concept of "employability" refers to "transferable skills and qualifications that reinforce people's ability to take advantage of education and training opportunities presented to them in order to find

and obtain job, to progress in his expectative carriers into the evolution of technology and labor market conditions. It is well known that one of the most important functions of the University is to apply the Knowledge generated by its research and Teaches with his teaching. It is imperative that their graduates, in addition to having a training guaranteed by their academic excellence, are inserted and widely integrated in the labor market, in order to apply the knowledge acquired and achieve the social purposes for those who graduated on campus University. Relevance evaluations points the employability of university graduates, observing the influence that, in terms of professional prestige, quality of employment and generation of a social good, has had university training in this sector of work. With the promulgation in 1995 of the law on prevention of occupational risks , and its subsequent development is approved RD 39/1997 Regulation of Prevention Services , regulates the development of the tasks of OHPS in Different levels, emerging the new figure of the Higher Technician in OHPS, which was required a previous university degree. Generated situation a new niche labor market, for professionals with a specific title and recognized in the field of OHPS. Likewise, in Annex VI of the same Royal Decree, the training programs corresponding to the higher level functions of OHPS are established. The training may be delegated and recognized by the competent labor authority (Ministry of Labor or Counseling with transferred competences) in favor of certain educational centers that meet the established requirements. This new legislation obliges companies with more than 500 workers, companies with between 250 and 500 workers to carry out activities of special risk or, in the case of companies not included in the previous sections, as determined by the labor authority, at their disposal In its staff of senior technicians in OHPS and in addition, that among its functions, it included specific training in OHPS to all its workers, with the main objective of reducing the casualty. For this reason, in the late 1990s, the OHPS was transformed into one of the labor sectors with the greatest demand for university graduates without having relevance to the degree of origin. In the same way, a large number of OHPS training centers, authorized by the labor authorities of the different autonomous communities, emerged to meet the emerging needs of the labor market on this type of professionals.

A number of other studies are summarized by who point out the two main approaches: the degree-day method and simulation techniques. Most of the papers in that review study office buildings and homes. The authors find that the predicted warming will result in a reduced heating load and an increased cooling load. This translates into a reduction in energy use for colder climates and an increase in electricity consumption for warmer climates.

In addition to the degree-day method and simulation, other approaches have been used. One example is the paper by Schrock and Claridge [2] in which the authors use a simple regression model of the ambient temperature to investigate a supermarket's electricity use. The use of multiple linear regression analysis allows the inclusion of any desired variable. This technique is used by Lam et al. [9] who study office buildings in different climates in China. These researchers include 12 input variables covering building parameters, building loads and the HVAC system in their regression model and find that predictions largely agree with building software simulation. Another example is Chung et al. [3] who investigate the energy use intensity of supermarkets by means of such diverse variables as operational schedule, number of customers, lighting control, employee behavior and maintenance factors, but explicitly exclude outdoor climate. Multiple regression analysis to investigate timer settings, night cover effectiveness together with indoor and outdoor temperature and humidity on the electricity consumption of a supermarket. The more complex principal component analysis is used by Lam et al. in for office buildings. This technique allows the same flexibility as multiple regression analysis, but is not restricted by its underlying assumptions (for more details see 2.3 Multiple linear regression analysis).The improvement of the qualifications and the professional competences that the university contributes to the graduates in the area of OHPS presents like one of the essential components of the employability.

Competently, professionals who have three of the four specialties existing in Occupational Risk Prevention, without taking into account occupational medicine (Occupational Safety, Industrial Hygiene and Ergonomics and Occupational Psycho sociology), are the most in demand for The companies for reasons of evident economic saving and efficiency being able to carry out all the preventive work. This is an aspect that normally meets most of the university masters offered by providing their graduates with the competences of the three specialties. There are now important advances in the training, with an appearance of a high Number of technicians with extensive knowledge in all the specialties of this discipline, many of them coming from university masters

Under the auspices of the EHEA, the university is entrusted with the task of training higher education providers, with a transversal, generalist, but specific training for each area of the training cycle in the different levels of education dependent on Ministry of Education.

In addition, of the 6.000 surveys carried out, as a source, of students graduated from university masters from all over Spain, on different aspects of prevention professionals, the study was expanded with specific research on employability and quality of employment Preventive, developed from data extracted from the 1st survey carried out in Spain on higher level prevention by the General Council of Industrial Relations and Sciences of the Work of Spain (CGRICT). The data obtained were collected and ordered through a survey conducted structured in questions a graduates of the whole national territory. The method of data collection of field developed allows to deepen in the knowledge about the degree of contribution that the university formation received by an exerted on its employability in the different productive sectors of the Spanish market.

In statistical analysis, regression models are often used for estimating the future events or values[8]. Trend extraction and curve fitting methods are also used to estimate the future behavior of the time series and to fit the future data according to the trend. Regression analysis includes parametric methods

such as linear and logistic regression [7-9].. Non-parametric methodologies such as projection pursuit, additive models, multivariate adaptive regression etc. have also been applied to estimation and prediction problems [7].

The proposed model is capable of forecasting the weather conditions for a particular station using the data collected locally[7]. The data is processed to obtain some statistical indicators to extract the hidden information present in the time series. These statistical indicators, also called as features viz. moving average (MA), exponential moving average (EMA), rate of change (ROC), oscillator (OSC), moments (μ2, μ3 and μ4) and coefficients of skewness and kurtosis are calculated over certain periods. On the basis of correlation, features are chosen as inputs to the models. Regression equations are obtained for the parameter to be forecast, which is termed as target. The whole data set is divided in two parts, the first is used to obtain MLR equations and remaining is used for testing the model [8,10].. The power of MS Excel has been used to process the data and present the results in simply understandable form. allow knowing the current situation of the OHP Sand their employability.

METHODOLOGY

Multiple linear regression models

Regression explains the nature of relationship, i.e., the average probable change in one variable given by a certain amount of change in the other variable.

The general equation of regression of Y on X is

$$Y= \alpha+\beta X+e \quad (1)$$

This equation is known as the mathematical model for linear regression. As the special case the form Y= a + β X is called the deterministic model (Agarwal, 1991). Multiple Linear Regression equation, which describes the linear relationship with set of dependent variable Y, and k sets of independent variables X1, X2, X3……Xk is,

$$Y= \beta_o+ \beta_1 X_1 + \beta_2 X_2 +.....+ \beta_k X_k +e$$

where Y is predicting and X1, X2, X3…..Xk are the predictors. 'e' is the error which is distributed normally with zero mean and variance s 2, i.e. e~N(0, s_2).

To develop the proposed model, the parameters a, β_1, β_2……..$ß_k$ are estimated using the training sample sets. X_1, X_2, X_3…….X_k are among the extracted features mentioned earlier. The coefficient of determination, r^2, explains the extent by which the variation of dependent variable Y is being expressed by the independent variable X. It is obtained as

$$r^2 = \frac{Variance\ explaine}{Total\ Variance} \quad (2)$$

Variance explained2 r = Total Variance

A high r2 shows that there exists a linear relationship between the two variables. If r^2=1, it indicates the perfect relationship between the two variables.

Extractions and selection of features for the models

The prediction model performs better when some hidden features are presented which enhance its adapting capability. The features considered in this study are statistical indicators calculated for different time frames and explained as below:

- Moving Average (MA): It is calculated progressively as an average of N number data values over the certain period. Data set is represented by dt, dt-1, dt-2 ,……., d0, where dt is present and d0 is the first data value, the moving average with a sliding window of period N is

$$MA = \frac{d_t+d_t-1 +d_t-2 +.....+d_t-n}{Total\ Variance} \quad (3)$$

-Exponential Moving Average (EMA): It is defined as

$$EMA = a \times (MA-EMA_{t-1})+ EMA_{t-1} .... \quad (4)$$

where 'a' is called the smoothing constant having value 0<a<1.

- Oscillator (OSC): Oscillator is used to indicate the rising or trailing trend present in the time series. It is defined as difference of moving averages or exponential moving averages of two different periods.

$$OSC = MA_{N1} - MA_{N2} \quad (5)$$

$$OSC = EMA_{N1} - EMA_{N2} \quad (6)$$

where N1 and N2 are different periods and N1>N2.

-Rate of Change (ROC): It indicates the rate of change of the variable at present, as compared to the value of the variable at certain period back.

$$ROC = \frac{1-dt}{dt-a}\ 100 \quad (7)$$

(v) Moments: The rth moment of X, μr is given by

$$\mu_r = \frac{\sum(X-\bar{X})^r}{N}\ 100 \quad (8)$$

Where X is total data set and is the mean of N data variables. Second, third and fourth moments correspondingly to r = 2, 3 and 4 are considered.

-Beta Coefficient of Skewness

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad (9)$$

- Beta Coefficient of Kurtosis

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \quad (10)$$

However, in some cases the parameter to be forecast exhibits a strong dependence on other weather parameters. In such cases the model should include the features extracted from other weather parameters also. Therefore, to forecast each parameter, independent variables are decided to include all these features exhibiting strong trends established by them over the period of observation. To assess whether a specific feature is suitable for the model or not its correlation with the target is obtained. Therefore, different features may be formed.

Mathematical models equations

First figure shows the target pool output and prediction from MLR model for employability over 6.000 surveyed the employment rate in occupational hazard prevention services (OHPS).

Second one figure shows the output show the training and testing scatter diagrams. During testing about outputs.

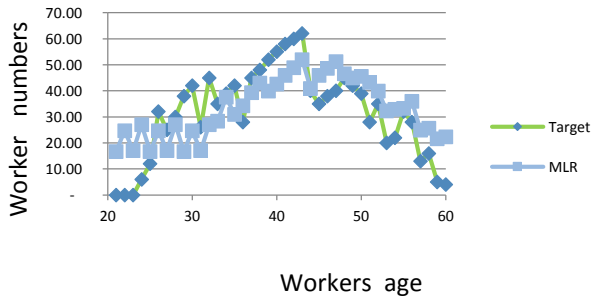

Fig 2:Trading target Vs output Engineering workers

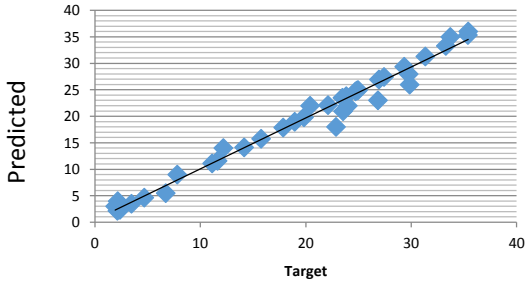

Fig 3:Scatter for estimating Engineering Workers

Table 2: Statistical regression analysis engineers polling

Statistical regression analysis	
Multiple Correlation Coefficient	0,9955
Coefficient of determination R^2	0,9910
R^2 adjusted	0,9899
Typical error	1,0531

RESULTS

Mathematical models equations

The data used are first used to obtain the regression equations that are shown in Table 1 and rest of the data are used to check the model. The regression equations are used to obtain the output parameters estimated by the regression model for training as well as for testing data. The scatter diagrams show the linear relationship between the actual weather parameter and the output of the MLR model.

Fig 4:Trading target Vs output Architecture workers

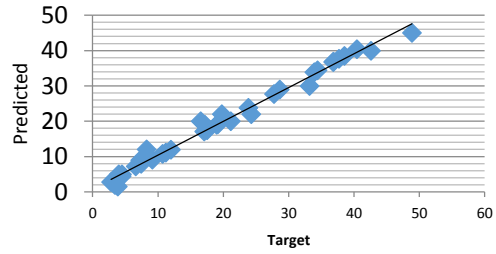

Fig 5:Scatter for estimating Architecture Workers

Table 3: Statistical regression analysis Architects polling

Statistical regression analysis	
Multiple Correlation Coefficient	0,9955
Coefficient of determination R^2	0,9904
R^2 adjusted	0,9892
Typical error	1,4094

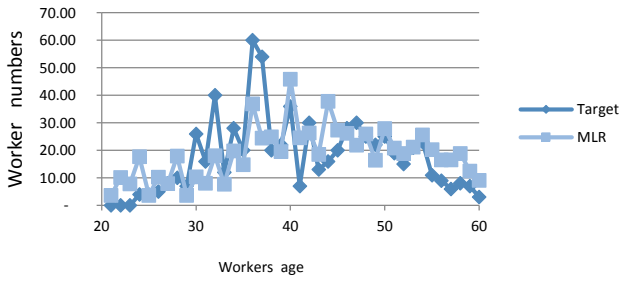

Fig 6:Trading target Vs output Socials Science workers

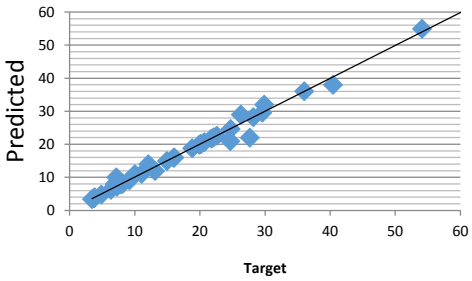

Fig 7:Scatter for estimating Socials Science Workers

Table 4: Statistical regression analysis Socials Science polling

Statistical regression analysis	
Multiple Correlation Coefficient	0,9985
Coefficient of determination R^2	0,9971
R^2 adjusted	0,9967
Typical error	1,0641

Fig 8:Trading target Vs output Medicine workers

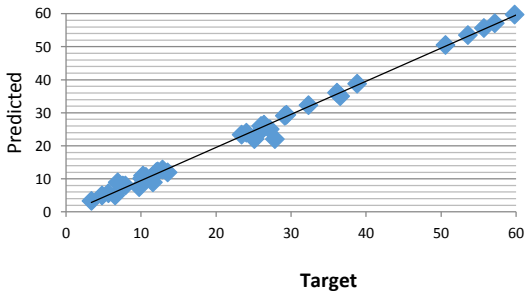

Fig 9 :Scatter for estimating Medicine Workers

Table 5: Statistical regression analysis Medicine polling

Statistical regression analysis	
Multiple Correlation Coefficient	0,9974
Coefficient of determination R^2	0,9948
R^2 adjusted	0,9942
Typical error	0.9297

The employment rate in the sector is 63, 58, the rates are distributed by Engineering (21.03%), Architecture (11.35%),

Social Science (11.83%), Medicine (16.18%), Humanist Science (11.40%). Le percentage de pooling workers is **71, 80**%, 4.308 employer over 6.000 polling workers

Coefficients of determination are average 0.995 with a minimal typical error 0.24 in Social Sciences and maximal 1.40 Architects that indicates quality regression to make a prognostic, **That indicates that the pooling respondents by the social sciences workers have not been precise, due basic training covers different areas of skills with a high probability to be overlapping his proficiencies with other works areas.**

CONCLUSION

The statistical forecasting models can be developed to estimate the future labor market in occupational hazard prevention services (OHPS).

In general, the results of the regressions ratifies the quality of the forecasts and give very important information to assessing the future of professional Spanish market in workers in occupational hazard prevention services (OHPS)

It is possible to relate the prediction avec separate parameter which correspond a male and female separate and get other order of range in multiple regression analysis

Technical and statistical indicators chosen are simple and capable of extracting the trends, which can be considered as features for developing the models. The model uses excel sheet data and Stat-graphics software to find the model parameters. Thus the method can be directly used by the medium educated farmers to forecast weather and other predicted aspects in engineering resource.

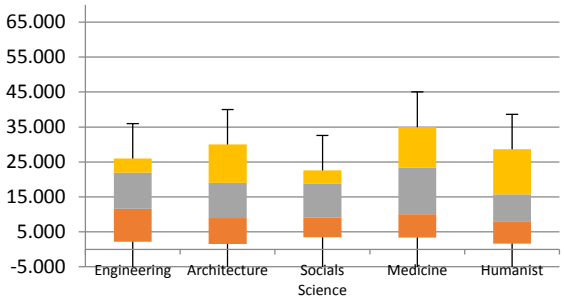

Fig12:Scatter error Polling rate Vs Predict Box Plots

References

[1] M.J.Azofra, (2000) "Cuadernos Metodológicos" Centro de Investigaciones Sociológicas. ISBN: 978-8474762723

[2] García Fernando M., (2003) "El análisis de la realidad social". Métodos y Técnicas de Investigación 3ª Ed.

[3] Segarra, A. y Callejón, M. (2002), "New firms' survival and market turbulence: new evidence from Spain", Review of Industrial Organization 20(1): 1-14.

[4] Agarwal, B.L., 1991, Basic Statistics, Second Edition, New Delhi, Wiley Eastern Ltd.

[5] Chatfield, C., 1994, The Analysis of Time Series- An Introduction, Fourth Edition, London, Chapman and Hall.

[6] Holmström, L., Koistinen, P., Laaksonen, J. and Oja, E., Neural and Statistical Classifiers: Taxonomy and Two Case Studies, Jan. 1997, IEEE Trans. Neural Networks, vol. 8, No. 8, pp 5-15.

[7] Montgomery, D. C. and Lynwood, A. J., 1996, Forecasting and Time Series Analysis, New York, McGraw-Hill.

[8] D.C. Montgomery, E.A. Peck, G.G. Vining Introduction to linear regression analysis John Wiley & Sons, Hoboken (2006)

[9] Mathur, S., Dec. 1998, Stock Market Forecasting Using Neural Networks-An MBA Project Report Submitted to School of Management Studies, IGNOU, New Delhi.

[10] Mathur, S., Shukla, A.K. and Pant, R.P., Agra, Sept. 22-23, 2001, International Conference on Optimization Techniques and its Applications in Engineering and Technology, A Comparative Study of Neural Network and Regression Models for Estimating Stock Prices.

[11] Roadknight, C.M., Balls, G.R., Mills, G.E. and Palmer-Brown, D., Modeling Complex Environmental Data, July 1997, IEEE Trans. Neural Networks, vol. 8, No. 4, pp852-861.

[12] Shukla, M.C., and Gulshan, S.S., 1980, Statistics, Fourth Edition, New Delhi, S.Chand and Company.

[13] J.C. Lam, K.K.W. Wan, D. Liu, C.L. TsangMultiple regression models for energy use in air-conditioned office buildings in different climatesEnergy Convers Manage, 51 (12) (2010), pp. 2692–2697

[14] M.G. Lawrence The relationship between relative humidity and the dew point temperature in moist air: a simple conversion and applications Bull Am Meteorolgical Soc, 86 (2005), pp. 225–233

[15] Park S.J., Lee Y.L., Numerical analysis on performance variation of linear compressor equipped with economizer. **International Journal of Applied Engineering Research** (2015)10(3) pp 5411-5420

[16] Ricardo Díaz Martín, GastonSanglier Contreras and Adib Guardiola Mouhaffel"Calculation of the Equilibrium Moisture Content of Pellets and Olive Stones using the Henderson Mathematical Model". **International Journal of Applied Engineering Research**, (2015) 12 (10) pp. 31681-31690