THE LABOUR MARKET IMPACT IN THE DOURO VITICULTURE: A FUZZY CLUSTER DISTRIBUTION APPROACH

Cátia SANTOS
Research fellow. University of Trás-os-Montes e Alto Douro (UTAD) and Centre for Transdisciplinary Development Studies (CETRAD)
cisantos@utad.pt

Aníbal GALINDRO
Research fellow. University of Trás-os-Montes e Alto Douro (UTAD) and Centre for Transdisciplinary Development Studies (CETRAD)
anibalg@utad.pt

Ana MARTA-COSTA
Corresponding author. Assistant Professor. University of Trás-os-Montes e Alto Douro (UTAD) and Centre for Transdisciplinary Development Studies (CETRAD), Portugal, www.utad.pt
amarta@utad.pt

Abstract

Demography and the labour market dynamics are recognized as a key factor for the development of a region. Performing labour market insights can be even more disjointed on mountain and steep slope regions where the hand labour is usually sparse and scarce. This study intends to find the labour influence into the core variables of a vineyard productive system such as production or overall revenues and realize their importance in the sustainability of the Douro region. Simultaneously, the sensibility of each variable upon external shocks was determined. The data was collected performing face-to-face inquiries directly from 50 Douro vine-farms and the Fuzzy Average with Fuzzy Cluster Distribution methodology was used for rank the variables. The results have shown that the labour costs seem to be the most influential variable on the vineyard production followed up by the steepness and terrain orientation. Nonetheless, the familiar labour also gains some prominence when the revenues are selected as the core variable. This study showed that the labour force plays an important role on the grape farm output and due to the current demographic trend in the region, the hand labour availability might be an alarming problem that threatens the sustainability of the Douro.

Keywords: Douro Region, fuzzy cluster distribution, labour influence, population dynamics, sustainable development

JEL classification: J43, Q12, Q19

1. Introduction

Agriculture can be characterized as a very heterogeneous sector, not only regarding the type of productive systems but also concerning the farm structure with expressive regional differences on farm size, topographic conditions, production and incomes (Bogdanov 2014). Simultaneously, there is a substantial contrast among regions in their demography and available hand labour. Within the agriculture umbrella lies the grape production towards wine making, an activity that gets fairly attention worldwide with, Italy, France and Spain capping the frontline of the world production (OIV, 2018). In Portugal, the wine business holds a large position of the agricultural sector, representing by the year 2016, 11% of added-value and providing 8% of employment to the population (INE, 2017).

Even though there were a lot of agricultural improvements during recent years due to the heavy mechanization and technological breakthroughs such activity still requires a fair amount hand labour. In fact, that requirement is even more pronounced in steep mountain viticulture when sometimes the task mechanization is substantially complex and troublesome which increases the hand labour requirements (Moreira and Guedes de Pinho, 2011). Such concern is very prominent in the Douro Demarcated Region (DDR), located in the Northeast of Portugal, where the steep slope vineyards are dominant.

It is also known that the labour market is closely related to the demographic dynamics and in this region there was a substantial reduction with the available labour force above 5%, between 2011 and 2016 (INE, 2017). Such fact is particularly explained with the exodus of the young people towards urban areas (Jan van der Laan, 2016) and also related to the agricultural low wages, leading the available workers to seek better paid activities that simultaneously increase their social prestige (Marta-Costa, 2010). Rebelo (2018) regards this trend as one of the major threats and challenges that the DDR will face in near future. Unfortunately, according to these authors, it is expected that such problem and trend might intensify in the next years with the expectation of a decaying active population of 2.8% for 2021. The lack of active population and following workforce unavailability should over-value the expected salaries or in an extreme scenario the total absence of labour supply and demand agreement. As a core and fundamental variable it is mandatory to address this problem as soon as possible.

Therefore, the labour shortage, nourished by the population dynamics can pose a real threat to the firm's productivity (Hossein et al., 2017). Bogdanov (2014) concluded that the demographic variations among regions may cause significant differences in productivity. So, the long-term depopulation had profound consequences on the agricultural production, leading to land abandonment, the decline of the overall land cultivation and diversification (Bogdanov et al., 2008).

Rebelo (2018) states as an utopic that the DDR wines need to see their overall prices augmented to economically solidify and address the actual concerns. This bold shift could surely allow the producers to pay higher wages and attract workers that otherwise would be unavailable. Other solution may lay on an analogous thought of Gibson and McKenzie (2014) and the New Zealand's Recognised Seasonal Employer (RSE) program. Such program aims to smooth and ease labour shortages on the horticulture and viticulture industries from New Zealand with migration programs from countries with severe unemployment. In this sense, focusing on the vineyards located in southern Italy, Seifert and Valente (2018) found an expressive increase of labour productivity with the 2011 migration wave, motivated by the employment of illegal workers with a significant decrease on the average hourly wages.

The wine sector can present itself as a crucial source of growth and development, contributing to stop the desertification in the rural areas, guaranteeing employment and better living conditions (Jan van der Laan, 2016).

This work aims to determine which are the variables of a grape productive system that have more effect on the farm`s production and revenues and, in this way, realize the real influence of the labour on the business sustainability. The sensibility of these variables upon external changes (such as the rural exodus or migration phenomena's) were also acquainted. To attain further information about the DDR demographic behaviour a time-bounded values were interpolated and a foreseen scenario for 2025 were obtained.

The present study provides the possible effects of labour shortage and the expected demographic changes, which may jeopardize the overall business sustainability. Besides, the results may be used for provide important political insights to solve the actual problem of the sector related with the sparse source of experienced labour. We also suggest an alternative mechanism to mitigate this problem based on temporary or circular migration programs and the labour supply from locations with lower living levels.

2. Data and Methodology

2.1. Surveys' data and methodology

The data was collected by face-to-face surveys of 50 vineyard grape growers in Douro Region applying the survey methodological framework according to Hill and Hill (2002).

The Douro Region has three sub-regions (Cima Corgo, Baixo Corgo and Douro Superior), four urban districts and 21 rural districts. The sample was selected from two rural districts of each sub region, realising a total of six rural districts. This choice was made considering the districts with larger representativeness regarding their total vineyards area. A quota (number of grape growers to be selected) was assembled to each district distributing them into four vineyard class areas to ensure de diversity and heterogeneity of the sample. The considered vineyard class areas were: a) $1 \leq$ area< 5; b) $5 \leq$ area< 10; c) $5 \leq$ area< 10 and d) ≥ 20 ha. The total sample size was 50, comprising of 16 surveys of the class area a); 17 of the class area b); 12 of the class area c); and 5 of the class area d). The reason why the first classes have a greater number of surveys is related to the small-scale land structure that is predominant in this region.

The variables used for this study are: Production (kg/ha); Revenues (€/ha); Labour costs (€/ha); Familiar Labour (%); Overall Costs (€/ha); Subsidies (€/ha); and the variables linked to the vineyard landscape: Steepness, Training systems which comprises guyot and cordon and Terrain orientation covering walled terraces (socalcos), patamares, vertical planting and plain vineyard. These variables were chosen based on their dominance on the mountain and steep slope Douro productive systems (Magalhães, 2015). Table 1 shows the average of the main variables for a general sample characterization.

Table 1. General information of the inquired sample

Variable	Values[1]
Area (ha)	9,75 ±1,39
Number of plots	4,97 ± 0,46
Production (kg/ha)	5571 ± 339
Revenues (€/ha)	3968 ± 257
Labour costs (€/ha)	3289 ± 287
Familiar Labour (€/ha)	1600 ± 293
Overall Costs (€/ha)	732 ± 64
Subsidies (€/ha)	3575 ± 427

Source: own elaboration

[1] Values are expressed as mean ± SEM (standard error of the mean)

2.2. Cluster methodology

Generally, the cluster analysis refers to a set of tools and methods, which try to subdivide a given dataset *X* into *y* subsets (called clusters). The core idea is to settle a function, which maps each individual point similarity to all the possible clusters (Zadeh, 1965).

The chosen cluster methodology for this study is the Fuzzy c-means clustering algorithm (FCM) developed by Ruspini (1970) and later extended by Dunn (1973) and Bezdek (1981). Lu et al. (2013) states that the FCM algorithm implements the clustering task for a data set by minimizing an objective-function subject to the probabilistic constraint, which accounts that the summation of all the membership degrees of every data point to all clusters must be one.

The classical FCM formulation suits the goals of this study more accurately since the partial membership assumption corroborates this article premise of meshed/ranked results per variable instead of a definitive cluster association typically applied on standard k-means methods. Nonetheless, the problem of this work is not purely an unsupervised problem, the present data structure compiles a well settled structure with dependent and independent variables. The attempts of Hou et al. (2007) and Lin et al. (1995) bundles the FCM approach with similar variable (input-output oriented) environment.

Similarly, to Hou et al. (2007), in this study there is a sample containing *m* data points with a single output variable (production or revenues) and *n* input variables (labour, steepness, terrain orientation, overall costs, training system, familiar labour, subsidies). The input and output data vectors are formed in Equations (1) and (2), respectively, where x_i^j x_i^j represents the *jth* measurement of the input variable x_i (with $i=1,2,\ldots,n$; $j=1,2,\ldots m$).

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \cdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} x_1^1 & \cdots & x_1^m \\ \vdots & \cdots & \vdots \\ x_n^1 & \cdots & x_n^m \end{bmatrix} \quad (1)$$

Similarly the output vector (2) acquaints that y^j represents the *jth* measurement of the output variable y_j (with $j=1,2,\ldots m$).

$$\mathbf{y} = [y^1 y^2 \cdots y^m] \quad (2)$$

The goal is to map the relationship between the input and the output variable such as the $x_i - y$ bundle is found (with i=1,2,..,n). The method of fuzzy curves used by Lin et al. (1995), originally in fuzzy-neural system modelling to settle the initial variable weights, accounts the definition of m fuzzy rules bonded to the sampled data (x_i^j, y_j). (x_i^j, y_j) Those fuzzy membership functions for each input variable x_i are Gaussian membership functions centered at x_i^j (Equation 3).

$$\mu_{ij}(x_i) = \exp\left(-\left(\frac{x_i - \bar{x}_i^j}{\sigma}\right)^2\right) \quad (3)$$

While $\bar{x}_i^j$ refers to the center of the membership function, σ σ accounts for the function's width. The result from the produced defuzzification is given in Equation (4).

$$C_i(x_i) = \frac{\sum_{j=1}^{m} y^j \mu_{ij}(x_i)}{\sum_{j=1}^{m} \mu_{ij}(x_i)} \quad (4)$$

The fuzzy-curves are calculated and the variables ranked according to their range. In order to acquaint the sensibility upon external changes, the Jackknife resample method (Rizzo, 2008) is performed in order to obtain 50 new simulations per variable. This method was performed leaving one observation out each time (49 observation) displaying the influence inner variability on the new simulations considering the variable range (maximum influence value minus the minimal influence value). The Jackknife attempts deliver the information about how the featured variable may vary upon new observations. The resampling method should shift the obtained curve (Equation 4) withstanding the information about each variable sensibility upon external shocks. Therefore, the interpretation should rely on the overall obtained range of the newly created 50 fuzzy curves.

With the variables above mentioned, three separated simulations were performed in the open-source R software (version 3.4.1). The selected inputs and output variables for the three simulations are available in Table 2 and the featured R code on Appendix 1.

Table 2: Description of the used variables and their range for the clustering analysis

Variable	Range	Variable description
Production	kg/ha	Average production of each farm per hectare
Revenues	€/ha	Farm overall income
Labour costs	€/ha	Costs with labour per hectare
Familiar Labour	0 to 1	Percentage (%) of familiar labour that each farm owns comparing to their own total labour force
Overall Costs	€/ha	Combines the value spent on electricity, gasoline, taxes, fertilizers and crop protection/pesticides
Subsidies	€/ha	Amount of financial aid that each farm yields
Steepness	1 to 4	Average steepness of each farm
Training system	0 to 1	Value 0 is the % of guyot training system that each farm owns. Value 1 is the percentage of cordon system
Terrain orientation	0 to 2	Value 0 is the plan vineyard landscaping; value 1 the walled terraces and *patamares*; and value 2 the vertical planting

Since the selected variables x_{in} with i=1,2,…9 and m=1,2,…,50 (regarding the 50 observations) acquaint such sparse numerical values, before applying the aforementioned method, they were normalized x_{im}^{n} between 0 and 1 according to Equation 5.

$$x_{im}^{n} = \frac{x_{im} - \min\{x_{im}\}}{\max\{x_{im}\} - \min\{x_{im}\}} \quad (5)$$

3. Results and discussion

Targeting the main goal of gathering information about the labour influence into the core variables of a vineyard productive system such as their production or overall revenues two simulations were made acquainting the dependent variable Production (kg/ha) and Revenues (euro/ha) respectively, while the independent variables remain the same for both simulations. The variables are ordered considering their overall influence on the dependent variable.

The first simulation (Table 3) delivers the labour costs (euro/ha) as the most influential variable from our data-frame followed up by the landscape variables. The straightforward interpretation lays on the fact that those variables are the most influential on the production of the featured 50 farm from the Douro region. In this region, the percentage of the workforce on the total operational costs varies between 70.6% in the mechanized vineyard (patamares) and 93% in traditional non-mechanized vines (Rebelo, 2018). This shows the importance of the labour factor undermined by the rational mechanization of the farm, which should be followed to fulfil the constraints imposed by the registration of the Alto Douro Vinhateiro in the UNESCO Heritage List (Andresen and Rebelo, 2013).

Table 3: Results of simulation 1

Production Simulation			
Rank position	**Variable**	**Influence Level**	**Jackknife Deviation**
1	Labour costs	0.00897	0.00487
2	Steepness	0.00755	0.00165
3	Terrain Orientation	0.00608	0.00083
4	Overall Costs	0.00466	0.00238
5	Training System	0.00375	0.00125
6	Familiar Labour	0.00367	0.00051
7	Subsidies	0.00268	0.00886

Source: own elaboration

The scenario shifts when the dependent variable is the farm revenue (Table 2), since the landscape variables switch places with the labour related one, nonetheless the remaining variables such as the overall costs, subsidies and the chosen training system remain with a peripheral influence on the featured dependent variable. This result is not farfetched at all since it displays the main characteristics highlighted previously on this article (labour intensive and mountain viticulture) that actually burdens the DRD farm production and following revenues.

The family labour is widely used in DDR, due to the large number of small farms, which in turn do not attribute any charge to this type of work. However, considering that the farms usually benefit from both familiar labour and public financial supports, such statement might be a reasonable explanation for the survival of a large number of small farms on this reason (Rebelo, 2018).

The influence levels calculated through the Jackknife methodology are also presented. It is possible observe that in both simulations (Table 3 and 4) the subsidies cap a substantial high value. Therefore, even though that variable doesn't seem to gather that much influence on the 50 featured farms further economic policies (such as lack or cancellation of the current subsidies) may scramble the actual scenario leading the subsidies to influence more positively or negatively the farmer's production/performance. This result corroborates the increasingly dependency on subsidies by the Portuguese agricultural sector with the accession to the European Community fund (Lains, 2016).

Table 4: Results of simulation 2

Revenues Simulation			
Rank position	**Variable**	**Influence Level**	**Jackknife Deviation**
1	Steepness	0.00950	0.00440
2	Terrain Orientation	0.00718	0.00161
3	Familiar Labour	0.00614	0.00208
4	Labour costs	0.00437	0.00152
5	Overall Costs	0.00237	0.00086
6	Training System	0.00211	0.00023
7	Subsidies	0.00149	0.00383

Source: own elaboration

An alternative to the referred dependency is the increase of the selling price of the grapes/wine. Some authors suggested that this bold shift is mandatory to solidify economically the Douro explorations. This audacious step could certainly allow the producers to pay higher wages and get workers easily. In addition, the rise in the price of the wine could lead to an higher income which may allow the farmers to be able to take more risks, as an example the investments in new technologies and innovation (Jan van der Laan, 2016; Rebelo, 2018).

This article highlights the labour force as the main research feature; therefore, we are interested in studying the working environment attributes that actually influences more or less the labour costs. A third simulation was performed, considering the labour costs (euro/ha) as the dependent variable featuring the steepness, the terrain orientation and the selected training system as the independent variable.

The results presented on Table 5 display the steepness as the most influential variable followed by the terrain orientation; the explanation may lay on the fact that those variables can seriously overvalue the labour costs related with the increased task difficulty. The training system remains itself as a substantially innocuous variable over the three simulations highlighting the fact that the decision between guyot and cordon may not be a concern upon the production and revenue levels and labour costs. Nonetheless, this variable may acquire a more meaningful interpretation on a farmer's decision in an expanded timeframe considering the vineyard's lifespan. The Jackknife results also entangle high variability to the most influential variables (the steepness and terrain orientation).

Table 5: Results of simulation 3

Labour Simulation			
Rank position	**Variable**	**Influence Level**	**Jackknife Deviation**
1	Steepness	0.04067	0.00799
2	Terrain Orientation	0.02275	0.00883
3	Training system	0.00608	0.00069

Source: own elaboration

The linear needs of labour consumption for Douro winegrowing are estimated at approximately 14,700 work units per year (AWU) differently distributed in a very asymmetric way throughout the year. For example, a daily requirement of about 19,500 people is estimated for 22 days only for the grape harvesting. Due to the low population density of the Douro wine-growing region and the seasonality of the activity, the labour needs traditionally have been supplied using temporary employment agencies that gather individuals from bordering territories (Rebelo et al., 2018).

With the purpose to attain further information about the Douro demography, Marta-Costa (2018) presented an hypothetical scenario for 2021. In 2011, the Douro had a total resident population of nearly 204 thousand inhabitants, approximately 6% of the North's total but equivalent to a comparing percentage loss of around 7.2 % in to 2001. The downward trend continues in this decade, with the population declining to around 193,000 in 2016. For 2021, Douro will have a population of around 183,000 people in 2025, representing an aggregated decrease of 11% compared to 2011.

In order to attenuate and reverse this trend, political insights are mandatory to create economic opportunities to ensure that the habitants are willing to live and remain in the Douro region.

Lastly, it is recommended that the public organisms encourage the farmer's education and awareness in order to introduce a new sustainability programme inspired by a particular case of the New Zealand's Recognised Seasonal Employer (RSE) program that intend to ease labour shortages, with migration programs from countries with severe unemployment (Gibson and McKenzie, 2014). In this sense, Seifert and Valente (2018) found an expressive increase of labour productivity in Italy with the 2011 migration wave.

4. Conclusions

This article provided a brief introductory endeavour regarding the Portuguese Douro region agricultural landscape, more specifically their grape-growers and wine producers. Supported by the available bibliography we have identified the labour force availability and cost as a key variable to the grape productive system. The FCM methodology was selected in order to rank the labour influence upon other farm variables. The results showed that the labour force plays an important role on the generalized farm output such as their overall production or final revenues. Alongside the previous statement, it was also found that the landscape variables such as the terrain orientation and steepness may also severally influence the output and also the general labour costs. The conclusion gains more relevance when money tangible variables such as the subsidies and overall farm costs (electricity, gas, water) acquire peripheral influence on the farm output. Even though it is well known that the Douro region is haunted by the mountain viticulture increasing difficulties the result overwhelms the fact that those features alongside the labour cost may play a core and important role in this region further survival. The demographic changes and increasing labour costs are quite alarming on such an influential variable.

Acknowledgements

This work was supported by the R&D Project INNOVINE & WINE – Vineyard and Wine Innovation Platform - Operation NORTE -01-0145-FEDER-000038, co-funded by the European and Structural Investment Funds (FEDER) and by Norte 2020 (Programa Operacional Regional do Norte 2014/2020).

References

Andresen, T., and Rebelo J., 2013. "Avaliação do Estado de Conservação do Bem Alto Douro Vinhateiro – Paisagem Cultural Evolutiva Viva" – Relatório de Avaliação. Porto.

Bezdek, J.C. 1981. "Pattern Recognition with Fuzzy Objective Function Algoritms." Plenum Press, New York, NY. DOI:10.1137/1025116

Bogdanov, N., Meredith, D., and Efstratoglou S. 2008. "A Typology of Rural Areas of Serbia", Economic annals. 177, 7-29. DOI:10.2298/EKA08177007B

Bogdanov, N.L. 2014. "The development of support for less favoured areas and deprivileged regions: challenge of agricultural policy in Serbia". Western Balk. Agric. Eur. Integr. EAAE 2014 Congr. Agri-Food Rural Innov. Heal. Soc. https://ageconsearch.umn.edu/record/186679

Dunn, J.C. 1973. "A Fuzzy Relative of the Isodata Process and Its Use in Detecting Compact Well-Separated Clusters". Journal of Cybernetics. 3, 32-57. DOI:10.1080/01969727308546046

Hill, M.M., and Hill, A. 2002. Investigação por Questionário. Edições Sílabo Lda. Lisboa.

Hossein, K., Taylor, T., and Goodrum, P. 2017. "Analysis of the impact of craft labour availability on North American construction project productivity and schedule performance". Construction Management and Economics. 35, 368-380. DOI:10.1080/01446193.2017.1294257

Hou, Y., Zurada, J., Karkowski, W., Marras W.S., and Davis, K. 2007. "Identification of Key Variables Using Fuzzy Average With Fuzzy Cluster Distribution". IEEE Transactions on fuzzy systems. 15 (4), 673-685. DOI: 10.1109/TFUZZ.2006.889897

International Organisation of Vine and Wine (OIV), 2018. OIV Statistical Report on World Vitiviniculture: 2018 World Vitiviniculture Situation. Accessed January 2019, available at http://www.oiv.int/public/medias/6371/oiv-statistical-report-on-world-vitiviniculture-2018.pdf

Gibson, J., and McKenzie, D. 2014. "The development impact of a best practice seasonal worker policy". Review of Economics and Statistics. 96 (2), 229-243. DOI:10.1162/REST_a_00383

Jan van der Laan, R. 2016. “Identifying key factors influencing sustainable wine production – A case study on the wine sector in the Porto and Douro region”. Utrecht University, Utrecht, The Netherlands.

Lains, P. 2016. “Agriculture and Economic Development on the European Frontier, 1000- 2000”, in D. Freire and P. Lains (eds.), An Agrarian History of Portugal, 1000-2000: Economic Development on the European Frontier, Leiden: Brill.

Lin, Y. 1995. “A New Approach to Fuzzy-Neural System Modeling”. IEEE Transactions on fuzzy systems. 3 (2), 190-198. DOI: 10.1109/91.388173

Lu, Y., Ma, T., Yin, C., Xie, X., Tian, W. and Zhong, S. 2013. “Implementation of the Fuzzy C-Means Clustering Algorithm in Meteorological Data”. International Journal of Database Theory and Application. 6, (6), 1-18.DOI: 10.14257/ijdta.2013.6.6.01

Magalhães, N., 2015. Tratado de viticultura - A videira, a vinha e o terroir . Esfera poética. Lisboa.

Marta-Costa, A. A. 2010. “Application of Decision Support Methods for Sustainable Agrarian Systems”. New Medit. 9 (2), 42-49.

Marta-Costa, A., Santos, M., and Galindro, A. 2018. “Dinâmicas demográficas”. In Hogg T. and Rebelo J., Rumo estratégico para o setor dos vinhos do Douro e Porto- Estudos de base. UTAD, Vila Real.

Moreira, N. and Guedes de Pinho, P. 2011. “Port wine”. Jackson, R.S (Ed), Advances in Food and Nutrition Research. (1st ed., vol 63), Elsevier Inc., USA.

National Institute of Statistics, INE (Instituto Nacional de Estatistica), 2017. População residente (N.º) por Local de residência (NUTS – 2013), Sexo e Grupo etário; Estimativas anuais da população residente. Accessed at October 2017, available at https://www.ine.

Rebelo, J. 2018. “Rentabilidade das explorações vitícolas: a relevância do fator trabalho”. In Hogg T. and Rebelo J., Rumo estratégico para o setor dos vinhos do Douro e Porto- Estudos de base. UTAD, Vila Real.

Rizzo, M. 2008. Statistical Computing with R, Chapman and Hall,CRC, New York.

Ruspini, E.H. 1970. “Numerical methods for fuzzy clustering”. Journal of information sciences. 2 (3), 319-350. DOI:10.1016/S0020-0255(70)80056-1

Seifert, S., and Valente M. 2018. “An offer that you can`t refuse? Agrimafias and migrant Labour on vineyards in southern Italy”. Discussion papers of DIW Berlim 1735, DIW Berlim, German Institute for Economic research. DOI:10.2139/ssrn.3180586

Zadeh, L.A. 1965. “Fuzzy sets”. Information and control. 8 (3), 338–353. DOI:0.1016/S0019-9958(65)90241-X

Appendix.1

R code

Influence results

```
my_data <- read.table(file = "clipboard",sep = "\t", header=TRUE)
Y <- as.array(my_data[,1])
X <- as.matrix(my_data[,2:ncol(my_data)])
U1 <- matrix(0, nrow(X),ncol(X))
U2 <- c()
for(j in 1:ncol(X)){
  for(i in 1:nrow(X)){
  Rv <- (range(X[,j],na.rm = FALSE))
  R <- Rv[2]-Rv[1]
  U2[i] <- exp(-(X[i,j]-mean(X[,j]))/(R*0.20))
  }
  U1[,j] <- U2
}
C1 <- c()
C2 <- matrix(0, nrow(X),ncol(X))
for(j in 1:ncol(U1)){
for(i in 1:nrow(U1)){
  C1[i] <- (sum(Y[i]*U1[i,j])/sum(U1,j))
}
  C2[,j] <- C1
}
for(i in 1:ncol(C2)){
  print(max(C2[,i])-min(C2[,i]))
}
sort(X[,2], decreasing = FALSE)
plot(X[,2],C1)
```

Jackknife leave one out simulation

```
#jackknife
my_data <- read.table(file = "clipboard",sep = "\t", header=TRUE)
Y <- as.array(my_data[,1])
X <- as.matrix(my_data[,2:ncol(my_data)])
U1 <- matrix(0, nrow(X)-1,ncol(X))
U2 <- c()
Res <-matrix(0, nrow=ncol(X),ncol=50)
C3 <- c()
for(m in 1:50){
 Y1 <- Y[-m]
 X1 <- X[-m,]
 for(j in 1:ncol(X)){
  for(i in 1:nrow(X)-1){
   Rv <- (range(X1[,j],na.rm = FALSE))
   R <- Rv[2]-Rv[1]
   U2[i] <- exp(-(X1[i,j]-mean(X1[,j]))/(R*0.20))
  }
  U1[,j] <- U2
 }
 C1 <- c()
 C2 <- matrix(0, nrow(X)-1,ncol(X))
 for(j in 1:ncol(U1)){
  for(i in 1:nrow(U1)){
   C1[i] <- (sum(Y1[i]*U1[i,j])/sum(U1,j))
  }
  C2[,j] <- C1
 }
 for(i in 1:ncol(C2)){
  C3[i] <- (max(C2[,i])-min(C2[,i]))
 }
 Res[,m] <- C3
}
for(z in 1:nrow(Res)){
 print(max(Res[z,])-min(Res[z,]))
}
```