Urban Roads Noise Modeling by Using Regression Analysis

Bülent BOSTANCI and Abdurrahman GEYMEN, Turkey

Key words: Road Traffic Noise Map, Regression Analysis, Interpolation

SUMMARY

In a daily routine of city life, noise is one of the most influential factors on human health. Noise, which may be denoted as undesired sound may also lead to physiological and psychological diseases. Particularly, big cities are the sources of relatively several noises such as transportation, industry and human voices. Developments of the big cities in the world have accelerated after the World War II. Parallel to the developing cities, numbers of motor vehicles have increased significantly and relevant traffic issues have arisen due to inadequate road network.

It is understood that the vehicle traffic has increased 10 folds between the periods of 1980 and 2012, in Turkey according to the statistical data. One of the most important issues in the developed and developing countries is the traffic noise due to the increasing number of the vehicles. Road traffic noise sourced by the vehicles present variety according to the density of the traffic, kind of vehicles, braking status, road surface and the air stream created by the vehicle.

In order to determine and assess the road traffic noise sourced by the traffic density, it has been planned to measure the noise on the main street in front of the Erciyes University, Kayseri within 5 minutes intervals in the morning, evening and at night. Noise values have been measured within a period for 3 three weeks of time under the similar air conditions along the route approx. for 2 km and in horizontal segments for 100 m by using Delta Ohm Type 1 noise measuring equipment. Position of the each spot measured on the land has been transferred to the satellite map by determining using the GPS Global Positioning System. It has also been recorded the vehicle numbers and types passed during the measurement duration for 5 minutes. Noise values of the day, evening and night noise levels measured on the same spot have been translated into a single value that reflects the daily average noise.

In this study, the relationship among the noise values obtained, road traffic density and the distance to the road have been studied by using the regression analysis. Besides this, noise values obtained as the result of the studies have been assessed by using interpolation methods and a Geographic Information System aided road traffic noise model have been created.

Urban Roads Noise Modeling by Using Regression Analysis

Bülent BOSTANCI and Abdurrahman GEYMEN, Turkey

1. INTRODUCTION

Urban road noise, which is under the scope of transportation noise, has emerged by the road transportation developed parallel to the technology and therefore brought a noise pollution together with the rapidly developing unplanned urbanization. For that reason, traffic noise shares the greatest part among the noise sources in Turkey (Şahinkaya, 2005). Noise modeling on the roads has been studied by using the multiple regression method. To and his colleagues (2202), have modeled traffic noise in Hongkong, 2002 along the 19 main roads and showed light vehicle number and total traffic flow as a significant factor in the urban traffic noise. Likewise, in his study Manatakis, in Patras, Greece, 2001 has defined total traffic per hour, light vehicle number per hour, heavy vehicle number per hour and average speed as an effective factor in modelling the road traffic noise in urban areas. In Turkey, Özer and his colleagues (2009) have studied on Tokat city noise pollution evaluations, Gündoğdu and his colleagues (2005) have studied on assumptions related to traffic noise by using genetic algorithm, Yılmaz and Hocanlı (2006) have studied on noise mapping in Şanlıurfa by using geographic information system. In 2010, according to the Regulation on the Assessment and Management of the Environmental Noise issued in Turkey, it has been required to prepare noise maps on the roads used by more than six million vehicles a year latest until 30.06.2013.

Researchers in almost all of the science branches are anticipated and requested working rules of the systems based on the data in hand. For that purpose, it is inclined to abstract patterns that will explain the system. Such abstraction is generally described as the word of "Model". Model is framing information or thoughts related to any event under defined rules. Thus, an approach to the reality may be achieved by getting more information related to the reality. However, information loss between the reality and model under the best circumstances is inevitable which means, observational errors in measurement and assignment due to uncertainties sourced by randomness is inevitable (Saygı and Şanslı, 2001).

Modeling art is trying to choose the most important factor by acknowledging the impossibility of consideration numerous single factors that have effect on the relevant variance instead. Later, a model which shows the interaction of such factors should be structured. Purpose is to reach a simple model that remarks workable model without ignoring the crucial factors (Newbold, 2000). Gujarati (1999) enumerates criterions of modeling such as (1) model should simple, (2) optimum balance of the model should be well, (3) model should be consistent with the theory, (4) Model's prediction should be high (Çağdaş, 2007). Regression analysis examines the dependency of a dependant variant to other significant variants in order to estimate the mean value (universe) of the former by the known or invariant values of the latter (in the repeated examples). Focus of interest of the regression analysis is statistical relationships and not functional or specific relationships among the variables seen in the conventional physics. Usually, random or stochastic variables, which are the variables that

have probability distribution, are used in the statistical relationships among the variables (Gujarati, 1999). Using estimation models, in other words, the regression analysis comprises the identification of other variables which are related to the variable whose values to be estimated. Statistical model formed after the variables identified defines the relationship between the variable to be estimated and the other variables and used for making prediction about the relevant variable. Reason of the intensive use of the regression models in the economy is that it allows evaluating the results of the optional policies of the management. It should be remembered that the difficulty in developing a model as well as the need for all the past data related to the whole variables have various disadvantages due to the need of time and cost (Bhattacharya, 1997; Özdemir vd., 2006).

Multilinear regression model hypotheses are as follows: (Kalaycı, 2008):

- Normal dispersion
- Linearity
- Mean of error terms are null
- Fixed variance
- Autocorrelation is null
- Multicollinearity among the independent variables is null.

2. MULTIBLE REGRESSION MODEL

In the relationship of a dependent Y variable with k number of variables such as X1, X2,……, Xk, if the independent variables are given the certain values of x1, x2,….,xk, multiple universe regression shows the Yi which is the value of dependent variable as follows:

$$Y_i = \beta_1 + \beta_2 . x_{1i} + \beta_3 . x_{2i} + ... + \beta_k . x_{ki} + u_i \qquad \text{i = 1,2,3,…..,n} \qquad (1)$$

Where β_1 constant term, $\beta_2, \beta_3, ...\beta_k$ are partial regression coefficients, u_i residue or error term and n is sample size.

Explanation power of the multiple regression model

If all the observations are on the regression line then a full concordance may be achieved. Such case may be seen rarely in the practice. Regression model certainly shall have some values with negative sign or $\hat{u}_i$ sign. It is anticipated that such errors should be minor as much as possible. R^2, the coefficient of specificity is a norm that how the sample regression line is in concordance with the data (Gujarati, 1999).

$$R^2 = 1 - \frac{\sum \hat{u}_i^{\,2}}{\sum (Y_i - \overline{Y})^2} \qquad (2)$$

Calculation of R^2 the coefficient of specificity with the matrix notation is shown with the formula (3).

$$R^2 = \frac{\hat{\beta}'.X'.y - n.\overline{Y}^2}{y'.y - n.\overline{Y}^2} \quad (3)$$

Coefficient of specificity describes the rate of change which may be explained by the relationship between the dependant and independent variables in the regression equation (Frees, 1996). Independent variables are used as a descriptive statistic which gives the success rate of the action of dependant variables. In this context, the use of k, the independent variable may be criticized when found in high rates in n, the observation point numbers. In this case, even if the dependant variable is not connected to independent variables strongly, it accords with the model data quite well. In order to alleviate such problem, a revised measure of regression accord power is used. Basic idea is to divide the total of both squares into a suitable independence degree in order to balance lowering total of square errors caused by adding more independent variable to the regression equation. By means of such change, the revised specificity coefficient is found (Newbold, 2000).

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \quad (4)$$

$\overline{R}^2$, revised specificity coefficient shows the rate of the changes in the dependent variable may be explained by the independent variables in % and valued between $0 \le \overline{R}^2 \le 1$. Values close to 1 show the high rate of the accordance optimization of the regression model. Coefficient of variation and mean percentile error are also used as accordance optimization measures.

3. APPLICATION OF MULTIPLE REGRESSION TO NOISE MODELLING

3.1. Work area and material

In order to determine and assess the road traffic noise sourced by the traffic density, it has been planned to measure the noise on the main street in front of the Erciyes University, Kayseri within 5 minutes intervals in the morning, evening and at night. Noise values have been measured within a period for 3 three weeks of time under the similar air conditions along the route approx. for 2 km and in horizontal segments for 100 m by using Delta Ohm Type 1 noise measuring equipment (Figure 1). Position of the each spot measured on the land has been transferred to the satellite map by determining using the GPS. It has also been recorded the vehicle numbers and types passed during the measurement duration for 5 minutes. It has been seen that the vehicles travel with a speed for 50 km in accordance with the inner city traffic. Therefore, mean speed value has not been calculated as independent variable. Noise values of the day, evening and night noise levels measured on the same spot have been translated into a single value that reflects the daily average noise (Lgag).

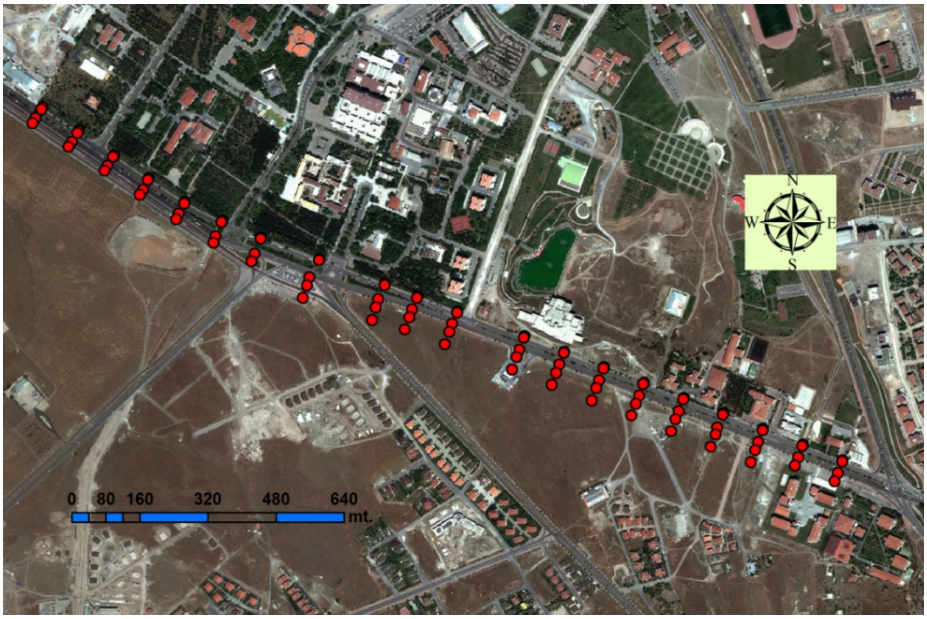

Figure 1. Satellite map and noise points of the work area

3.2. Multiple regression used to road noise modelling

It has been studied on roads noise modeling by using multiple regression method. To and colleagues (2202), have modeled traffic noise in Hongkong, 2002 along the 19 main roads and showed light vehicle number and total traffic flow as a significant factor in the urban traffic noise. Likewise, in his study Manatakis, in Patras, Greece, 2001 has defined total traffic per hour, light vehicle number per hour, heavy vehicle number per hour and average speed as an effective factor in modeling the road traffic noise in urban areas.
Today, multiple regression analysis, which is one of the statistical analyses, may be applied through the information technologies effortlessly is a kind of analysis related to predicting dependent variable based on two or more independent variables (expressive variables) connected to dependent variable. Linear multiple regression analysis observes to assess the independency of dependent variable to independent variables, mean value of dependent variable (universe) and independent variables in terms of known or unchanged values (Gujarati, 1999; Çağdaş, 2007)

$$Y_i = \beta_1 + \beta_2 . x_{1i} + \beta_3 . x_{2i} + ... + \beta_k . x_{ki} + u_i \qquad (5)$$

In the equation, Yi is the noise value wanted, x_i shows the variables that may possibly affect such value. Coefficients that define the efficiencies of the variables affect the noise value are the error terms of the model that has been defined as the mean to be null in $\beta_1, \beta_2, \beta_3, ..., \beta_k$ and u_i' Application of the multiple regression analysis in the assessment of noise is shown in Figure 2 in details.

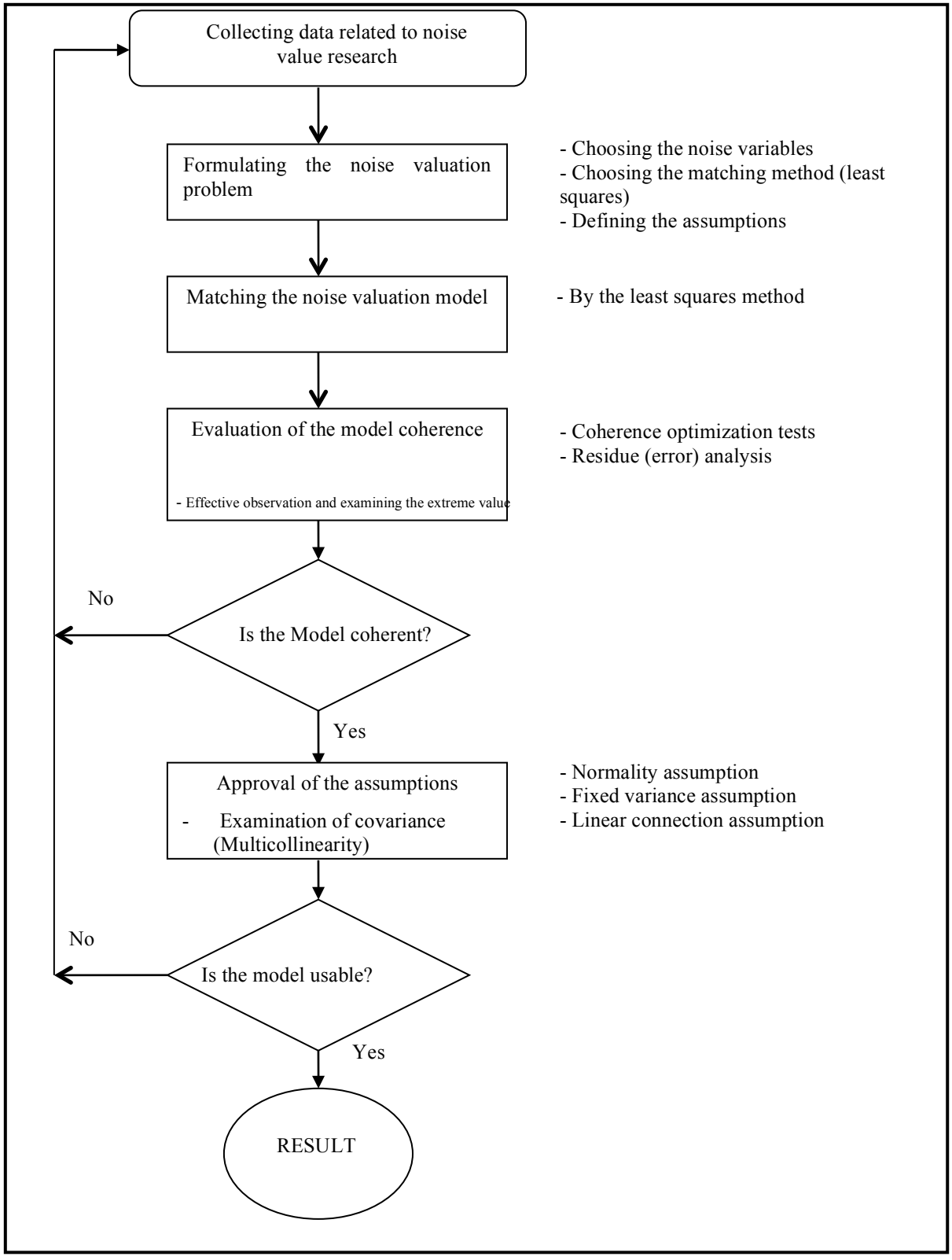

Figure 2. Flow chart of linear multiple regression analysis used for assessing the noise

Data obtained by the SPSS program has been assessed by multiple regression analysis and stepwise regression model has been used as the method of choosing variables (Table 1). Mean noise value obtained by the measurements of day, evening and night for 5 minutes, instead of dependant variable, and 4 variables, such as the distance of the noise measurement points to the road, total vehicle number, light vehicle number and heavy vehicle number within 15 minutes as independent variables have been used.

Table 1. Mean and Standard deviations of the variables

	Mean	Std. Deviation	N
Lgag	75,9890	3,20787	94
neardist	17,07	16,764	94
Total vehicles	261,81	22,417	94
Light vehicles	251,74	21,026	94
Heavy vehicles	10,06	4,706	94

It has been observed a correlation between the total vehicle numbers and lightweight vehicle numbers such as 0.97 which is significantly high when correlation coefficients are examined. Model shall prefer one of such variables. Any extreme values which will affect the model in the error statistics have not been observed (Table 2).

Table 2. Statistical results of the residues

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	69,5534	79,7141	75,9890	2,54719	94
Std. Predicted Value	-2,527	1,462	,000	1,000	94
Standard Error of Predicted Value	,223	,774	,394	,111	94
Adjusted Predicted Value	69,3655	79,4263	75,9818	2,55088	94
Residual	-4,72182	3,86864	,00000	1,94995	94
Std. Residual	-2,382	1,952	,000	,984	94
Stud. Residual	-2,480	2,099	,002	1,010	94
Deleted Residual	-5,11810	4,47361	,00719	2,05521	94
Stud. Deleted Residual	-2,555	2,140	,001	1,018	94
Mahal. Distance	,185	13,199	2,968	2,402	94
Cook's Distance	,000	,172	,014	,027	94
Centered Leverage Value	,002	,142	,032	,026	94

$\overline{R}^2$ Value, revised specificity coefficient, which is the cohesiveness degree of the model, has been found as 0,618. Rate of the loss in the model is 38,2%. Standard error of the assessment has been found as 1, 98 dBA. As the Durbin Watson coefficient, which is used to express the autocorrelation has been found as 1, 48, it may be reported that a autocorrelation is not existed (Table 3).

Table 3. Cohesiveness degree of the model

Model	R	R Square	Adjusted R Square	Std. Error of the Estimation	Durbin-Watson
1	,757a	,574	,569	2,10595	
2	,783b	,613	,605	2,01687	
3	,794c	,631	,618	1,98218	1,480

As the F test shows null significance in the Anova table used for testing the variance analysis and regression equation, regression model created is significant statistically (Table 4).

Table 4 Analysis of Variance (ANOVA)

Model		Sum of Squares	df	Mean Square	F	Sig.
3	Regression	603,399	3	201,133	51,192	0,0000
	Residual	353,613	90	3,929		
	Total	957,012	93			

In order to evaluate the effect of the each independent variable, it must be referred to the regression coefficient table of the output (Table 5).

Table 5 Coefficients of the variables and fixed value

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
3	(Constant)	71,849	2,556		28,111	0,000
	neardist	-0,134	0,013	-0,703	-10,688	0,000
	Heavy vehicles	0,122	0,045	0,179	2,714	0,008
	Light vehicles	0,021	0,010	0,136	2,053	0,043

Regression equation:

It has been formed as follows:

$$L_{gag} = 71,849 + (-0,134)x\,neardist + 0,122\,x\,heavyvehicles + 0,021\,x\,lightvehicles$$

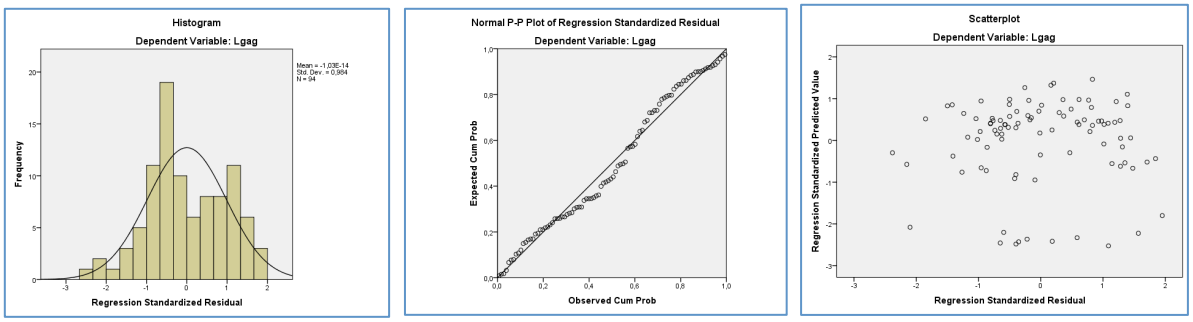

Figure 3 Histogram, dispersion and scatter graphics of the errors

Standardized error terms frequency graphics for the dependent variable must show a normal dispersion. Histogram graphics show that such requirement has been met. It may be said that standardized error values dispersion graphics describe a linear relationship and points tend to meet around a linear axis. In a scatter graphics, observations on dependant variable and standardized error terms must scatter homogenously on a defined area. It may be said that homogenous dispersion have been met partially (Figure 3).

3.3. Mapping road noise

Researches made in various countries show that the most significant and common source of the noise that affect numerous people in a residential area is the traffic noise (Berglund, vd., 1995; Akerlöf, 1996; Maraş v.d., 2011; Ilgar, 2012). Furthermore, researches have shown that road traffic noise is lot more significant than the noise sourced by the air traffic (Karpuzcu, 2007). Source of the noise of the road traffic may be described as vehicle types, engine parts of vehicles, noise of the vehicles generated by their own speed, quality of the road and the intensity of the traffic (Öztürk, 1999). According to the Regulation on the Assessment and Management of the Environmental Noise, strategic noise maps shall be prepared showing the status of the previous calendar year on the roads used by more than six million vehicles a year until 30.06.2013 (Bostancı ve Geymen, 2012).

Location and noise values found have been assessed by means of geostatistical analysis module of the ArcGIS software and a noise mapping model has been created through the interpolation techniques based on GIS. Kriging method is used and the noise map created by Ordinary Kriging Method is shown in Figure 4.

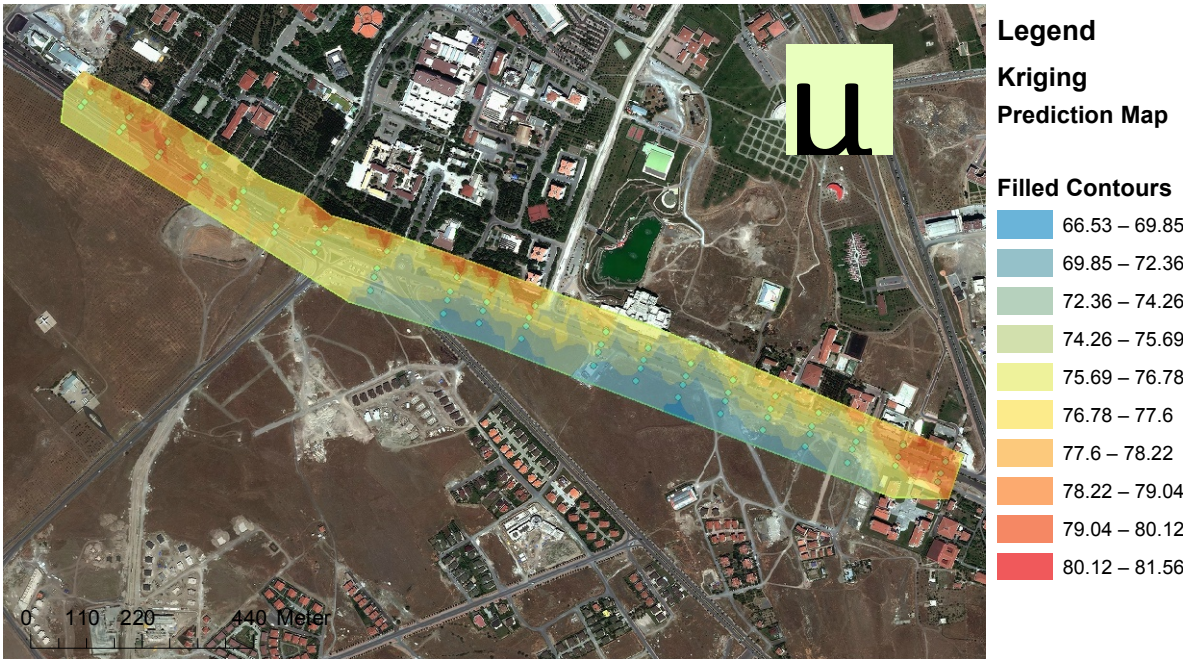

Figure 4. Urban road noise map by using Ordinary Kriging method

4. CONCLUSION AND RECOMMENDATIONS

Model meets all hypotheses. However, in order to see the results of the application of the noise regression sample which is found on other samples, it is necessary to apply on data separated out of the model for testing purposes.

A modeling is made with 4 variables that explain the dependant variable in the sample and a model has been created as the stepwise regression model considered 3 variables significant.

Cohesiveness optimization of the regression model increases up to 61.8% which is a low rate. Such low rate issue may be solved by increasing the observation times to 10 minutes, lowering the distances between the sequences to 50 meters and increasing the number of the variables.

REFERENCE

Akerlöf, L. (1996). Acoustic Guide, Selection of Acoustic Quality in Buildings. Sweden.

Berglund, B., Lindvall, T., 1995, Community noise. Report to the World Health Organization. Archives of the Center for Sensory Research., 2:1-195.

Bhattacharya, S., (1997). A Comparative Study of Different Methods of Predicting Time Series, M.Sc. thesis, Concordia University, Canada.

Çağdaş, V., (2007), Türkiye İçin Bir Emlak Vergi Sistemi Tasarım Modeli Önerisi, Doktora Tezi, Yıldız Teknik Üniversitesi, İstanbul.

Çevresel Gürültü Eylem Planı, (2009-2020),

Frees, E. W., (1996), Data Analysis Using Regression Models: The Business Perspective, Prentice-Hall: New Jersey.

Geymen, A., Bostancı B. (2012). Production of geographic information system aided noise maps. FIG Working Week 2012, Rome, Italy.

Gujarati, D.N., (1999), Temel Ekonometri, (Çev., Ü.Şenesen, G.G. Şenesen) Literatür Yayınevi, İstanbul.

Gundogdu, O., Gokdag, M., Yuksel, F.: A traffic noise prediction method based on vehicle composition using genetic algorithms, Applied Acoustics 6 (2005), pp. 799-809.

http://www.cygm.gov.tr/CYGM/Files/EylemPlan/Cevresel_%20Gurultu_20Eylem_Plani.pdf

Ilgar, R., (2012). Çanakkale şehir içi trafiğindeki araç kaynaklı gürültü kirliliğine yönelik ön çalışma, Journal of World of Turks, 4,1, 253-267

Kalaycı, Ş., (2008), SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil yayın Dağıtım Ltd.Şti., Ankara.

Karpuzcu, M., (2007). Çevre Kirlenmesi ve Kontrolü, 380 sh., Özal Matbaası, İstanbul.

Manatakis E.K., (2011), A multiple Regression Approach for Statistical Modeling of Road Traffic Noise, Proceedings of the 12 th International Conference on Environmental Science and Technology, Rhodes, Greece. 8-10 September 2011.

Maraş E.E., Maraş H.H., Maraş S.S. ve Alkış Z., (2011). CBS Verilerinden Çevresel Gürültü Haritalarının Hazırlanmasında Kullanılan Tahmin Yönteminin Analizi, Harita Dergisi, 145, 52-60

Newbold, P., (2000), İşletme ve İktisat İçin İstatistik, (Çev.,Ü.Şenesen) Literatür Yayıncılık, 4. Baskı, İstanbul.

Ozer, S.; Yilmaz, H.; Yesil, M., Yesil, P.: Evaluation of noise pollution caused by vehicles in the city of Tokat, Turkey, Scientific Research and Essay, 4 (2009), pp. 1205-1212.

Şahinkaya, S. (2005). Coğrafi Bilgi Sistemleri ile demiryolu gürültü kirliliğinin modellenmesi: Konya örneği. Selçuk Üniversitesi, Fen Bilimler Enstitüsü, Konya

Saygı, H. ve Şanslı, Ş., (2001), "Su Ürünleri Tüketimi İçin Ekonometrik Bir Model", E.Ü. Su Ürünleri Dergisi, Cilt 18, sayı 3-4, Syf. 383-390, İzmir.

To W.M.,, Ip R.C., Lam G.C., Yau CT. (2002) A multiple regression model for urban traffic noise in Hong Kong. Journal Acoust. Soc. Am., 112(2), 551-556,

Yilmaz, G., Hocanli, Y.: Mapping of Noise by using GIS in Sanliurfa, Environmental Monitoring and Assessment 121 (2006), pp. 103-108.

Özdemir, A. ve Özdemir, A., (2006), Ege Akademik Bakış, Ekonomi, İşletme, Uluslararası İlişkiler ve Siyaset Bilimleri Dergisi, Cilt 6, Sayı 2, Temmuz, Syf. 105-114.

BIOGRAPHICAL NOTES

Bülent Bostancı was born in Aksehir, Turkey, in 1969. He graduated from Yıldız Technical University, Istanbul, Turkey in 1992, and received MSc and PhD degrees from Yıldız Technical University, Istanbul, Turkey in 1995 and 2008. He worked at Republic of Turkey General Directorate of Highways for 5 years. He received his Assistant Professor degree in Turkey in 2010. He is now working for Erciyes University, Kayseri, Turkey as the Head of the Surveying Technical Division at the Department of Geomatics Engineering. His research interests are special engineering surveying, real estate development and risk analysis.

Abdurrahman Geymen was born in Konya, Turkey, in 1973. He graduated with a MS degree in Geomatics Engineering from the Gebze Institute of Technology, Turkey, in 2000. He received a PhD degree from the Department of Geomatics Engineering, Karadeniz Technical University, Turkey, in 2006. He is working with the Department of Geomatics Engineering, Erciyes University as an associate professor. His research interests include Geographical Information System (GIS), Computer-aided Design, Software Development for GIS, Web based GIS and Spatial Data Structures. He has published a number of international journal and conferences article related to these areas.

CONTACTS

Assoc. Prof. Dr. Abdurrahman Geymen and Assist. Prof. Dr. Bülent Bostancı
Erciyes University, Faculty of Engineering, Department of Geomatics, 38039 Kayseri/TURKEY
Tel. +90 352 2076666 (ext. 32676 and 32655)
Fax +90 352 4375784
Email: ageymen@erciyes.edu.tr and bbostanci@erciyes.edu.tr
Web site: http://harita.erciyes.edu.tr/en/page/16/academic-staff.html