DEVELOPING A MODEL TO PREDICT PEDESTRIAN MOVEMENT IN URBAN SPACES BY INCORPORATING SPACE SYNTAX AND EPR;
CASE STUDY KHAZANEH NEIGHBORHOOD IN CITY OF TEHRAN-IRAN

Mostafa Abbaszadegan [a], Hossein Babapoor [b]

[a] Iran University of Science and Technology

[b] Master of urban planning, University of Science and Technology

Corresponding author: Abbaszadegan@iust.ac.ir

©Ontario International Development Agency ISSN: 1923-6654 (print)

ISSN 1923-6662 (online). Available at http://www.ssrn.com/link/OIDA-Intl-Journal-Sustainable-Dev.html

Abstract: Predicting intensity and various uses of urban open spaces is an important urban open space management issue. Also, it is an important challenge for urban planners and designers to pursue their goals in urban environment. In a research that examines the use of urban spaces in zone 16 of city of Tehran, the capital city of Iran a predictive model has been developed using Space Syntax method and Evolutionary Polynomial Regression. An encounter model has been used to survey the pedestrian movement in the city.

Several research experiments indicated that space syntax technique can develop models that can interpret and predict pedestrian movement in regard to urban morphology, regardless of this capability it has some shortcoming, that is not able to specify the effect of all dependent and independent variables and their combination as a factor on pedestrian movements. To overcome this shortcoming Evolutionary Polynomial Regression has been utilized.

Keywords: Space Syntax, and Evolutionary Polynomial Regression, pedestrian, urban space, Tehran

INTRODUCTION

Urban open spaces play an important role in fulfilling the need for face to face contact which is an essential issue in socio – psychological wellbeing as well as formation of sense of community (Gehl 2010). The main objective of urban design is to enhance quality of urban environment and improve the citizen inter-relationship with it. To achieve these vital objectives in urban design process, a certain scenario should be programmed that enhances the relationship of people and environment within the urban context.

One of the main concerns of urban planners and designers as well as urban managers is to be certain that which scenario and what aspect of urban open space qualities can increase presence of citizen and enhance quality of their relationship in urban spaces.

In this research, the type of relationship of variables that have effect on level of presence of people on space has been explored in order to develop a predictive model. To develop the above predictive model, spatial value of morphological indices that has been obtained through space syntax analysis, land use variables and number of pedestrians in each space has been used as a fundamental data to analyze the relationship of urban morphology components and level and type of spatial use.

To conduct this research, Depth-map software was manipulated to extract spatial property of the neighborhood using space syntax method. A GIS data bank assisted to obtain number of shops that located along each axial line of the neighborhood. Finally number of pedestrian has been recorded in public spaces of the neighborhood via encounter observation method; then GIS software was used to make an analytical layer by superimposing three layers of the above information. To have effective variables in the model, correlation analysis of variable assessed and then EPR software (this technique has been introduced by Giustolisi, Savic & Doglioni, 2004 for inventing "Evolutionary Polynomial Regression" according to generic programming "EPR"). EPR assisted to select variables in a cyclical repetitive process and eventually a proper model produced and then MATLAB model presented the process.

THE MAIN ISSUE

In the past, urban spaces were mainly created as artistic phenomena, therefore many of the predicted uses by the artist planners were never realized and the created spaces were not beneficial ones. However, there were always few successful artistic designs, due to the deep vision of the artists that appreciated the requirements of the society. As a result of a rapid population growth and urban expansion, extensive urban development has occurred. The complexity of urban relations and lack of adequate knowledge about the rapidly changing social structure complicated the prediction of urban events in urban spaces; this resulted in failure of most artistic predictions.

Many attempts have been undertaken to explain and predict people's behavior and its relation to physical fabric, but none of these attempts were a comprehensive method which can explain a direct relation between urban physical morphology and behavior. Such a research need resulted in invention of the space syntax method in late 70's and its further development during 80's and 90's (Hillier & Hanson, 1984). Now this technique is widely used to analyze the urban pattern throughout the design process in many countries. Space syntax method is used to analysis various events and their relationship with spatial properties of urban morphology. One the type of research that this article is base on it is relationship of movement and spatial properties.

The most important and effective factors in promoting movement in the city are the origin and destinations. The study of movement generation is considering these two major factors and fails to interpret the distribution of movement flows in the distribution network that connects the origin and destination points (Hillier 1993). Distribution networks are the routes which are not necessarily the origin or destination of the pedestrians, but are used to move through to get from origin to destination. The mechanism that one selects a movement routes within the distribution network is well explained by Space Syntax method and techniques. Based on Space Syntax theory, the selection of the intermediate route to get to the destination is directly related to the syntax of urban spaces.

The relation between the structure of cities and the movement density is called the "natural movement". As Hillier explains, natural movement is part of the movement which is due to the structure of urban space that he called it "spatial configuration" – and not the attractions of attractors within the city. Various researches have shown that much of the movement within the city (in any space) is not due to the existence of origin or destination focal points in that space, but due to its presence on the route from origin to destination. It must be noted that even the planned movement (from origin to destination) has to move through intermediary spaces. Many studies – e.g. London, Barns Berry (Hillier, 1996, 1997) showed that three quarter of the movements in any urban space is due to the spatial configuration. In other words, three quarter of pedestrians' movement within any urban space is there to pass through and only a quarter of the pedestrian's presence in any space is due to the existence of origin or destination in that space.

It is important to mention that not only the spatial property of space defines the global structure of movement, but also the local characteristics of space would encourage people to stay longer in the space and benefit from its positive characteristics. When people are passing through one space due to its spatial characteristics, the land uses required by the population are concentrated in such spaces and in turn it produces movement. The results of various researches confirm that there are strong correlations between the space syntax characteristics and the use of spaces, density of movement, types of land uses and the adjacent land values (Abbaszadegan, 1999; Turner, 2007; Hillier, 1996).

The results of studies conducted by Hillier et al, 1993; Hanson, 1993; Pepoinis, 1989) and also the studies conducted in the space syntax Laboratory in University College London, show that the level of global integration of spaces could have a strong correlation with level of pedestrian movement and concentration of retail stores. Also a study was conducted in Iran using the same method by Abbaszadegan (1998) with comparable results. In this study, in addition to the three parameters of integration, number of retail stores and the number of pedestrians, other parameters presented in Table 1 were used to conduct the analysis.

The improvements of computer techniques and improved processing, capacity and speed of processors have created a new pioneer for researchers to analyze the natural phenomena through experimental data gathering. Among these techniques, the semi- accidental technique has gained an increasing validity among other techniques. Evolutionary Algorithms are expanding continuously with new applications and their potential of analysis is increased. One of these methods is the Evolutionary Polynomial Regression. This is a recently developed technique; it was first used to explain the environmental phenomenon for its users. (Giustolisi, Savic & Doglioni, 2004; Giustolisi, and Savic, 2006; Giustolisi and Berardi, 2007; Giustolisi, Doglioni, Savic and Webb, 2007).

Later this model was used in other areas including civil engineering (Javadi and Rezania, 2009). While natural phenomena are ordered and systematic, still there are tremendous complexities and ambiguities, which make their study and analysis very difficult and uncertain. The study of pedestrian flow and their behavior in urban space is one of these natural complex phenomena which its prediction relies on many unknown parameters. This makes prediction very difficult and uncertain. As noted, due to progress in computer techniques, now there are techniques available for analysis and forecasting of such phenomena, one of this is Evolutionary Polynomial Regression. This is conducted by EPR software.

The high speed and capability of this software in modeling process is one of its main characteristics and advantages. This research which was looking for a technique to predict pedestrian flow in urban spaces is a pioneering research using this technique.

The high speed and capability of this software enables the modeling processes to be conducted based on trial and error, and many tests could be done in a very short period of time.

A research was conducted in zone 16 of Tehran Municipality _ Khazaneh area, and the result of that is being reported in this article. The study uses Space Syntax method, encounter observation and Evolutionary Polynomial Regression technique to develop a predictive model for pedestrian flow in urban spaces.

RESEARCH METHODOLOGY

In this paper in order to find and develop a model to explain the systematic relation between pedestrian flow and presence in urban spaces with physical characteristics of urban fabric, the following procedure is used. (a) Spatial and physical analysis of zone 16 of Tehran (KHAZANEH District) using Space Syntax technique (b) Observation and enumeration of the pedestrian flow in urban spaces using the encounter observation technique in the sample area (c) Completing the database with the number of retail units along each axial line within the study area (d) Examining the correlation of all spatial parameters (shown in table 1) with pedestrian flow (e) Modeling using the Evolutionary Polynomial Regression Technique

Space Syntax

Space Syntax method examines and analyses the relation of all urban spaces and shows the results as graphically shown mathematical parameters. The graphic presentation of Space Syntax is very useful tools for urban studies, in which the results of physical interventions in urban morphology could be presented in graphic form. Space Syntax is very successful in predicting pedestrian flow as well as level of space use.

Space Syntax is a set of theories and techniques regarding the phenomenology of space. This theory is mainly based on philosophy and mathematics and its unique computer software is modeling urban spaces and graphically illustrates them. It predicts the results of designs and decisions for users and uses of space. One of the most important applications of space syntax is analysis of pedestrian flow within urban system. By utilizing space syntax, designers and planners can predict movements before the real development occurs within the city.

The Space Syntax research tries to have fresh look at the built environment as an integrated spatial system and analyze it through its spatial syntax and sequence of spaces and their relations. Space Syntax explains the way urban physical syntax could be analyzed through controllable factors and variables, so that the overall system could be analyzed through these variables. Through physical analysis of urban mass, the physical structure of the city and its relation with present functions can be analyzed.

Space Syntax helps to recognize the consequences of changes in urban form, especially changes in street networks on users perception and hence citizen's behavior. For this, the syntax of all urban spaces in a sequential structure is analyzed (Hillier, 1996).

It could be concluded that Space Syntax is a set of theories and techniques that analyze spatial morphology with particular application in architecture and urbanism. This is both a spatial model for recognition as well as a practical quantitative method for analyzing urban structures and patterns.

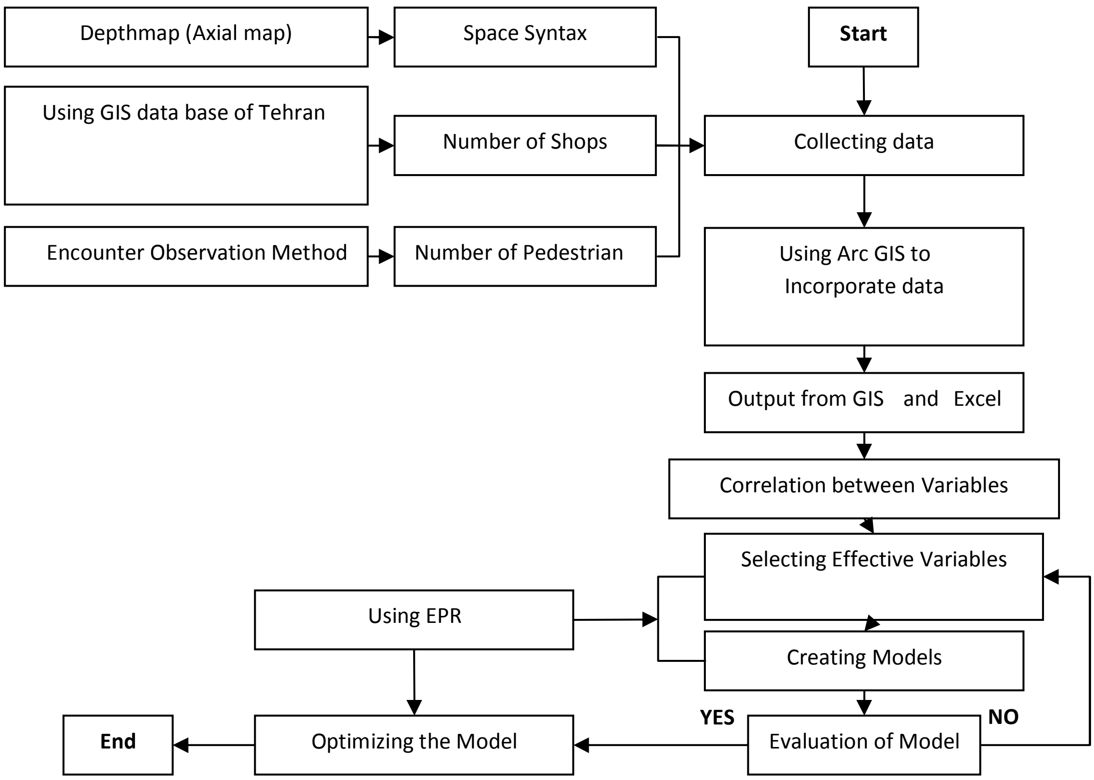

Figure 1: Research process

Evolutionary Polynomial Regression

Davidson et al (1999, 2000) developed a new regression method known as Rule-Based Symbolic Regression (R-BSR) to make polynomial models based on both quantitative and symbolic regressions (Davidson J.W., Savic and Walters 1999; Davidson, Savic and Walters, 2000). Giustolisi et al (2004) introduced the EPR model based on this model. Similar to R-BSR, EPR method is also a two phase technique for making symbolic models. This includes making the structure and estimating the parameters. The major differences of the two methods are in the search method and development of the preliminary structure of the model. In EPR method, instead of using the tree based GP of R-BSR, a simple genetic algorithm is used. In EPR first the search for symbolic structures are conducted by GA, and then constants are determined through a simple linear least square (LS). Following is a short explanation of EPR technique (Giustolisi, Savic & Doglioni, 2004; Giustolisi and Savic, 2006; Giustolisi and Berardi, 2007; Giustolisi, Doglioni, Savic and Webb, 2007).

At first the following equation is used.

$$Y_{N\times1}(\theta,Z)=\left[I_{N\times1}\quad Z^{j}_{N\times M}\right]\times\left[a_0\quad a_1\quad \cdots\quad a_m\right]^T=Z_{N\times d}\times\theta^T_{d\times1} \quad (1)$$

In this equation

$Y_{N\times1}(\theta,Z)$ is the least square estimate vector of N target values;

$\theta_{1\times d}$ is the vector of d=m+1 parameters aj, j = 1:m, and a0;

$Z_{N\times d}$ is a matrix formed by I, unitary column vector for bias a0, and m vectors of variables Zj that for a fixed j are a product of the independent predictor vectors of variables/inputs, $X=\langle X_1 X_2 \ldots X_k\rangle$

The matrix of inputs X is given as;

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1k} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2k} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3k} \\ \dots & \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & x_{N3} & \dots & x_{Nk} \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 & \dots & X_k \end{bmatrix}$$

(2)

Where the kth column of X represents the candidate variables for the jth term of Equation (3). Therefore, the jth term of Equation (1) could be written as;

$$\forall j = 1 \dots m$$
$$Z^j_{N\times 1} = \left[(X_1)^{ES(j,1)} \cdot (X_2)^{ES(j,2)} \cdot (X_3)^{ES(j,3)} \cdot \ldots \cdot (X_k)^{ES(j,k)} \right]$$

(3)

In this equation Zj is the jth column vector whose elements are products of candidate-independent inputs and ES is a matrix of exponents. Therefore, the problem is to find the matrix $ES_{k\times m}$ of exponents whose elements can assume values within user- defined bounds.

The original EPR have formulated according to single objective algorithm (SOGA) strategy for exploring the formulae space.

In 2006, Giustolisi and Savic developed EPR further as they used "Multi – Objective Genetic Algorithm" (MOGA) strategy. The main two objectives of this method was: (1) maximization of the model accuracy and (2) minimization of the number of polynomial coefficients in the formulae.

In 2007 Giustolisi & Berardi initiated another improvement in the use of EPR for predicting pipe failures; (1) the improved model called the Multi-Case EPR Strategy (MCS-EPR). This model could develop a comprehensive model. Typical flow diagram for the EPR procedure is shown in figure 2.

The case study

The present study is conducted in zone 16 of Tehran, (Khazaneh district). Figure 3 shows the location of zone 16 of the greater Tehran. This is middle to lower class district in terms of socio-economic status. It has a high residential density of 500 persons/hectare, and is one of the relatively dilapidated areas of Tehran. This is due to its access system that transformed to an almost isolated area. It is encircled by Bethat highway from north, railway from west and Bethat power station from east. Its only proper connections are from south. Therefore, it is functioning independently and is not much affected by surrounding areas.

First, the axial map of the study area is prepared; this is a simplified graphic representation of streets and urban open spaces which is used as the basis for spatial analysis of sequence of urban spaces. The axial line is the longest visual and access line in an urban environment. Therefore the axial line includes the structure of a set of urban spaces which are developed based on the longest accessibility-visibility axial lines. Axial maps include all urban public spaces. (Hillier B. & J.Hanson, 1984)

In preparing the axial map of the study area, and in order to minimize the marginal errors in space syntax analysis, a larger area was selected for analysis. Figure 4 illustrates the integration map of axial lines in the study area.

Both global and local analysis was conducted, using a rotation radius of 3. Then, the volume of pedestrian flow was determined through observation. Also the number of commercial units along axial lines was needed. Due to time and budget limitations, a sampling method was used. There were 166 axial lines in the study area. Using chockran method, a sample size of 116 was needed to achieve a 95% confidence level. Surveying pedestrian flow volume, encounter observation technique has been employed.

First, in a pilot exercise, the surveyor moved in one of the most populated axes in a busy time of the day, in a steady pace in a manner which enabled him to count all the pedestrians. This exercise was used to determine the highest acceptable speed during survey. After setting the criteria for accepted speed of movement of observer during observation, the survey was conducted along the selected axes and counted the number of pedestrians.

Survey of commercial units

For this part of the study and in order to improve the speed and precision of data, GIS software was used. First the depth-map data was exported into GIS software. Then, GIS data bank of the local detailed plan was used to attach the number of commercial units along axial lines; the result was saved in an Excel file format. In order to complete the databank the pedestrian volume for each axial line was also inserted in other columns of the excel file. Then data analysis and selection of data for modeling was conducted.

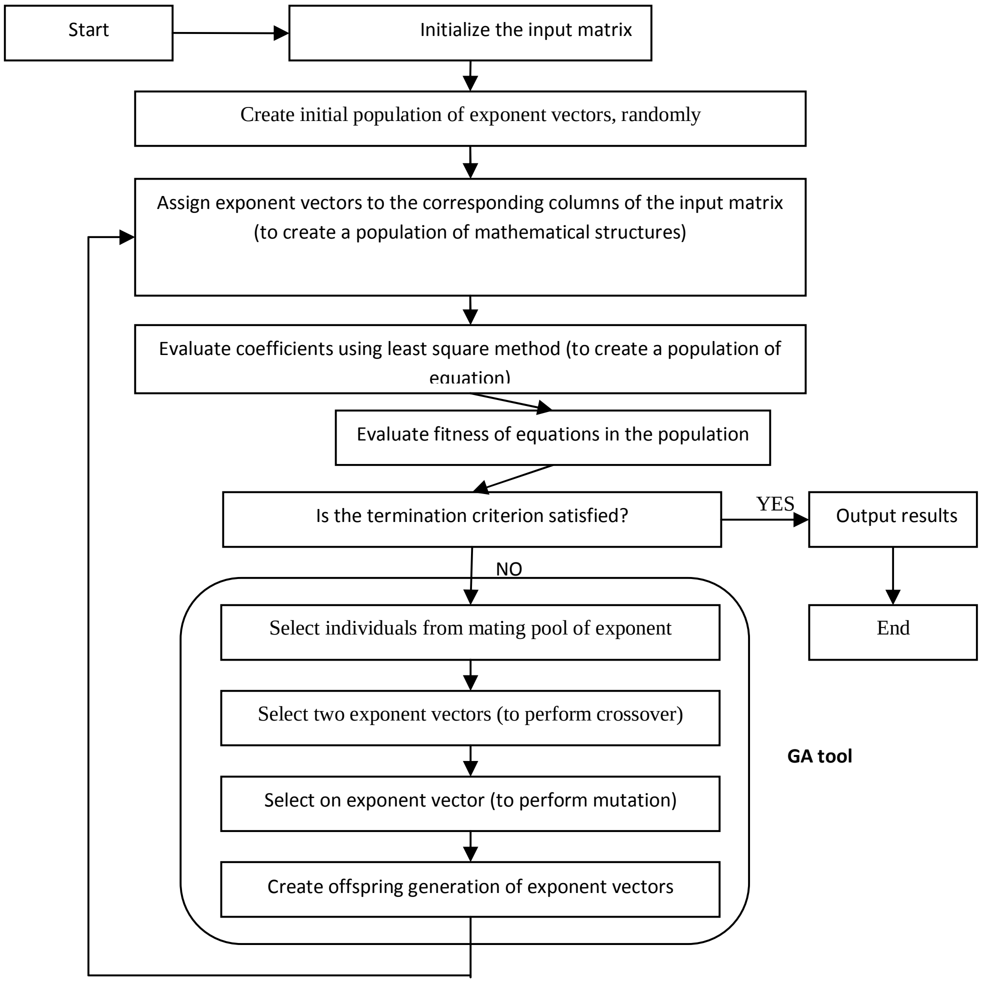

Figure 2: Typical flow diagram for the EPR procedure (Rezania, Javadi and Giustolisi, 2008)

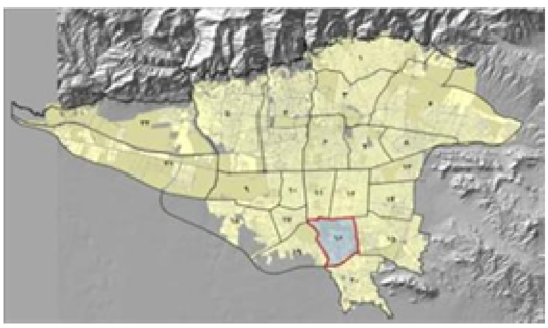

Figure 3: Zone 16 in the city Tehran

Figure 4a: Integration map of 16 zone of Tehran

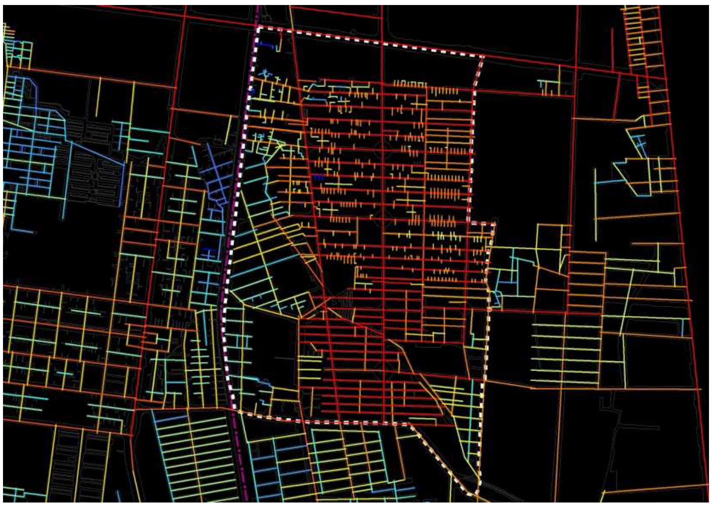

Figure 4b : Integration map of the study area (Khazaneh neighborhood)

Table 1: correlation coefficient of spatial parameter and volume of pedestrian movement in each axial line

ϖαριαβλε	ναμε	χορρελατιον(Ρ) (ς:Πεδεστριαν_Χουντ)
ϑ	Λινε_Λενγτη	0.884
Ρ	Στορε	0.850
Ε	Ηαρμονιχ_Μεαν_Δεπτη_Ρ3	0.692
Α	Χοννεχτιϖιτψ	0.670
Η	Ιντενσιτψ	0.517
Π	ιντ_ΗΗ	0.508
Θ	ιντ_ΗΗΡ3	0.497
Φ	Ιντεγρατιον_Τεκλ	0.467
Γ	Ιντεγρατιον_Τεκλ_Ρ3	0.464
Ι	Ιντενσιτψ_Ρ3	0.419
Μ	Νοδε_Χουντ_Ρ3	0.408
Χ	Εντροπψ_Ρ3	0.351
Δ	Ηαρμονιχ_Μεαν_Δεπτη	0.084
Β	Εντροπψ	–0.031
Ο	Ρελατιϖισεδ_Εντροπψ_Ρ3	–0.397
Κ	Μεαν_Δεπτη	–0.410
Ν	Ρελατιϖισεδ_Εντροπψ	–0.439
Λ	Μεαν_Δεπτη_Ρ3	–0.493
ς	Πεδεστριαν_Χουντ	1.000

DATA ANALYSIS

Correlation study

At this stage the relation and level of correlation between space syntax parameters (spatial properties) extracted from depth-map analysis and number of commercial units and number of pedestrians are studied. For this purpose Excel software is used and the results are reported in Table 1.

Table number 1 shows that correlation of spatial parameters of the variable J, R,E,A,H and P with pedestrian volume in each axial line are: .884,.850,.692,.670,.517 and .508.

Selection of data for modeling

It is important to find out the appropriate variables in modeling. For this purpose a trial and error procedure was conducted using EPR software. This enabled examining models in a comparable environment and assisted the selection of the best model. For this purpose, the data were selected for all possible conditions and the program was run. The results of the trials which were considerably reliable are as follow: (a) Selection of all data (b) Selection of all positively correlated data (c) Selection of all highly (0.5) correlated data (d) Selection of all relatively high (0.4) correlated data

The preliminary results of these four conditions are presented in table 2.

Table 2 shows that the best result are achieved using all positively correlated data.

Table 2: Result of the conditions of the selected data

Χονδιτιον	Παραμετερσ									ΧΟΔ	ΣΣΕ
1	A	B	X	Δ	E	Φ	Γ	H	I	98.048	42.659
	ϑ	K	Λ	M	N	O	Π	Θ	P	97.831*	49.257*
2	A		X	Δ	E	Φ	Γ	H	I	99.318	19.970
	ϑ			M			Π	Θ	P	97.688*	61.855*
3	A				E	Φ	Γ	H	I	99.147	22.341
	ϑ			M			Π	Θ	P	98.086*	52.377*
4	A				E			H		99.225	18.524
	ϑ						Π		P	97.211*	71.749*

The above results are without considering effect of variable R in model

Table 3: parameters in EPR method

Fields	Adjustment	Parameters
1	Y=sum(ai*X1*X2*f(X1)*f(X2))+a0	EPR Type
2	Static	Regression Type
3	No Function	function
4	21 statements (120)	Terms
5	(-2-1-0.5012)	Expon
6	No use in this research	na, nb + nk (Time Order)
7	No use in this research	Scale Input-output
8	LS	Scale Input-Output
9	15	Solution
10	Plot, Bias	Seed, Plot, Bias
12	-	Objective function
13	(min(aj, SSE))	MOGA Strategy

Table 4: Resulted models

Model number	Model
1	=3.9518e-014*(A)^-0.5*(E)^2*(C)*(G)^-2*(H)^-0.5*(I)^2*(J)^2*(M)*(P)^-2*(Q)*(R)+0.008816*(D)^2*(F)^-2*(G)^2*(H)^-1*(P)^-2*(Q)*(R)+0.0032752*(E)^2*(C)^-2*(F)^-2-0.055903*(A)^2*(C)^-2*(D)^2*(F)^2*(G)^2
2	=1.564e-013*(A)^-0.5*(F)^-1*(G)^-1*(H)^2*(I)^-1*(J)*(M)^2*(P)*(Q)^2*(R)^2+0.75497*(F)^-1*(H)^-1*(P)^-0.5*(Q)^-1*(R)+0.00095723*(F)*(G)*(I)^2*(M)^2*(P)^-1*(Q)
3	=2.9612*(P)^-2*(R)+0.029923*(E)^2-0.10649*(A)^2*(P)+1.2467e-012*(A)^2*(E)^2*(H)^-2*(J)^2
4	=0.025253*(A)^-0.5*(E)^2*(H)^2*(P)^2-0.47159*(E)^-0.5*(H)^2*(P)*(R)^2+0.028558*(H)^-2*(P)*(R)^2+1.3342*(P)^-0.5*(R)-0.0028528*(E)^2*(H)^-0.5*(J)^-1*(R)^2-4.5744e-006*(A)*(E)^2*(H)^-2*(R)-0.52672*(A)^2*(E)^-1*(H)^-2*(J)^-0.5*(P)^2*(R)+8.7932e-007*(A)^2*(E)*(P)^-2*(R)^2+3.593e-013*(A)^2*(E)^2*(H)^2*(J)^2*(P)
5	=0*(E)^-0.5*(F)^-2*(G)^2*(H)^2*(J)^2*(M)^-0.5*(P)^-1+0.0072634*(F)^2*(G)^-2*(H)^-1*(J)*(P)^-1*(Q)^2+9.7546e-014*(A)*(F)^-0.5*(G)^-2*(H)^-2*(I)^-1*(J)^2*(M)^2*(P)^-2*(Q)^2
6	=0.0024295*(A)^-1*(E)^2*(H)^-0.5*(J)*(P)^-2-0.010485*(A)^2*(H)^-2*(P)+1.5101e-013*(A)^2*(E)^2*(H)^2*(J)^2*(P)^2

Table 5: characteristics of each model

model	variable									Training			Testing			Rank
										COD	SSE	correlation (R)	COD_V	SSE_V	correlation (R)_V	
1	A		C	D	E	F	G	H	I	98.42	38.98	0.9982	83.12	35.65	0.866	4
	J			M			P	Q	R							
2	A		C	D	E	F	G	H	I	99.47	13.84	0.9993	83.9	31.9	0.915	2
	J			M			P	Q	R							
3	A				E			H		98.37	41.05	0.9980	84.42	34.61	0.932	1
	J						P		R							
4	A				E			H		99.2	20.59	0.9989	81.26	40.13	0.907	3
	J						P		R							
5	A		C	D	E	F	G	H	I	96.82	82.4	0.9964	63.2	82.61	0.869	2
	J			M			P	Q								
6	A				E			H		95.9	111.13	0.9950	72.49	59.93	0.935	1

However due to relatively similar results among conditions 2, 3 and 4, the process is not considered as a final stage. In order to evaluate the models, and examine their prediction precision, model was built using 70% data (training) and the extra 30% data was used for prediction model (Testing). These models have the ability to deal with a great number of information and to learn complex model functions from examples, i.e. by 'training' using sets of input and output data (Giustolisi, and Savic, 2006). In order to compare and assess models, COD, SSE, COD_V AND SSE_V parameters were used. The first two were for training and the second two were for testing models. Higher COD (COD _ V) and lower SSE (SSE_V) showed more reliable results.

Model accuracy, or fitness to observed data, is evaluated using the Coefficient of Determination (CoD) and SSE is stands for Sum of Squared Errors. The first two were for training and the second two were for testing models. Higher COD (COD_V) and lower SSE (SSE_V) showed better results. COD (COD_V): shows the Correlation between the real data and the data resulted through modeling. The higher numbers (closer to 100) shows better results.

SSE (SSE_V): The important issue about SSE is that this is a relative number. It is related to the size and quantity of the data. In other words, if the data was multiplied by a constant, SSE would also be multiplied by the square of that constant, while COD does not change.

MODELING USING EPR

EPR software is used for modeling. This software was developed based on programming in MATLAB. For modeling using EPR, there are specific parameters which help improve the modeling results. These are as follows.

EPR type

This popup menu defines the structure of the expression that you want to look for. The decision about the structure depends on your prior knowledge about the phenomenon you want to model. There are seven available model structures, see Giustolisi & Savic (2006), where the invertible function g assumes to be as logarithm, exponential, sinus and tangent functions.

Regression Type

In this popup menu you have to choose the kind of model you wish to construct. There are three available choices

Dynamical model

Modeling of a dynamical system, data are arranged as time series.

Static model

Modeling of a static system, data are not necessarily arranged according time series.

Classification

Modeling of a static system, the output of the model is required to be an integer value.

Type of Function

In this popup menu, you can choose the function type f(X), see Giustolisi and Savic (2006). You have five possible choices available, the function f(X) is chosen according to your prior knowledge about the phenomenon at stake. The five possible choices implemented so far are: no function, natural logarithm and exponential function, hyperbolic tangent and hyperbolic secant.

Terms

In this field, you have to assume the range for the number of terms constituting the potential set of models

Expon

In this dialog box, you have to set the values for the exponents. These are the values the elements from X are raised to.

na, nb + nk

This field is used if the menu n. 2 (Regression Type) is set as Dynamical. In this case you have to set the time order of the past measured.

Scale the input and output value

In this double dialog box you have to set the range in which you wish to scale the input values (left) and the output values (right). This scale option is suggested for the construction of dynamical models in order to increase the stability of the models.

Solution

In this popup menu, you can choose how the constant values aj are estimated. There are three possible choices: Least Square estimates, Linear Programming, Least Square aj > 0 (Giustolisi, Doglioni, Savic and Webb, 2007).

Gen

In this box, you have to set a proportionality parameter for the number of generation of the GA.

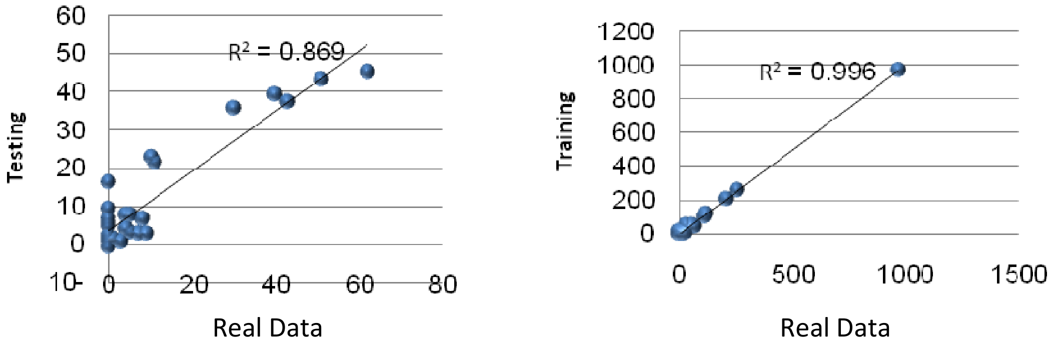

Figure 5: Correlation coefficient value of real data and predicted values of model number 3

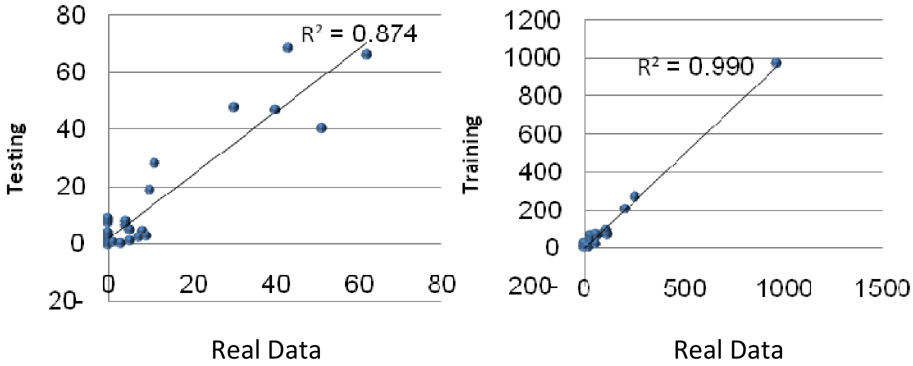

Figure 6: Correlation coefficient value of real data and predicted values of model number 6

Seed, Bias and Plot

Switching on the button of Seed, during its running, EPR seeds the population of formulas with random elements from the previous parental set. This option efficiently works when large data sets are available and in single-objective configuration.

Objective function

These radio buttons refer to the objective function to be optimized, which EPR uses for the structure construction. As it is mentioned, EPR uses as cost function the minimization of the sum of squared errors (SSE).

MOGA Strategy

This popup menu gives the option of selecting the type of gen algorithm of SOGA or MOGA.

Various test conducted by changing the parameters values for resulting an optimum model, the best result can be seen in table number 3.

RESULTS

The results of the study are presented through six models which can predict the number of pedestrians in urban spaces using variables in table 1. Table 4 shows the predictive models and Table 5 shows the characteristics of each of these models. The difference of model 5 with models 1 and 2 and model 6 with models 3 and 4 are in R variable.

To measure validity of models such items as correlation coefficient (R), SSE and COD are used. Table 5 shows that all six models are highly valid. As an example model 3 in Training mode has a correlation coefficient of 0.998, and in Testing mode has a correlation coefficient of 0.932. Also the validity of the models is shown in figures 5 and 6 in which a comparison of the predicted data and the real data for models 3 and 6 are illustrated

CONCLUSION

As explained the main objective of this research was to develop a model that be able to predict pedestrian flow in urban spaces. For this purpose the integration parameters and the number of shops along the street was used. Also other parameters of Spatial Value analysis that gained by the Space Syntax method were used to improve the predictability of pedestrian flow. The model was developed using EPR software. The remarkable capability and speed of this software enables modeling based on trial which is conducted in a very short period of time. The results of these trials show that some of the Parameters as presented in Table 1 contribute to the predictability of the model.

Considering the validity of the model and its capability to predict pedestrian flow in an axial line providing the physical parameters, the model could be used for: (a) Designing a vital urban space (b) Decision making for locations of subway and other public transport stations (c) Prioritization of servicing urban spaces, e.g. at the start, Predicting quantity and placement of service and amenities. (d) Determining appropriate width of pedestrian walkways (e) Increasing or decreasing the intensity of use of a street through changes in network (routes)

REFERENCES

[1] Abbaszadegan M;(1999), the relationship of urban spaces and users behavior, in: Benson J, Urban Life style, Balkema, Landon.

[2] Datta M., Gulhati S.K. and Rao G.V. (1982), "Engineering behavior of carbonate soils of India and some observations on classification of such soils", STP 777, ASTM, pp. 113-140.

[3] Davidson J.W., Savic D.A. and Walters G. (1999), "Method for Identification of explicit polynomial formulae for the friction in turbulent pipe flow", Journal of Hydroinformatics 2:1, pp. 115-126.

[4] Davidson J.W., Savic D.A. and Walters G. (2000), "Approximators for the Colebrook-White formula obtained through a hybrid regression method", Computational Methods in Water Resources vol. 2, pp. 983-989.

[5] Gehl Jan, 2010, Cities for People, Islan Press, Washington

[6] Giustolisi O. and Berardi L. (2007), "Pipe level burst prediction using EPR and MCS-EPR", Proc. of Computer and Control in Water Industry (CCWI) - Water Management Challenges in Global Changes, Ulaniki et al. (eds), Taylor & Francis Group, London, pp. 39-46.

[7] Giustolisi O., Doglioni A., Savic D.A. and Webb B. (2007), "A multi-model approach to analysis of environmental phenomena", Environmental Modeling and Software, 22:5, pp. 674-682.

[8] Giustolisi O., Savic D.A. & Doglioni A. (2004), "Data Reconstruction and Forecasting by Evolutionary Polynomial Regression" 6th Int. Conf. on Hydroinformatics, Singapore, Liong, Phoon & Babovic (eds), World Scientific Publishing Company, vol.2, pp. 1245-1252.

[9] Giustolisi, O. and Savic, D.A. (2006), "A Symbolic Data-driven Technique Based on Evolutionary Polynomial Regression", Journal of Hydroinformatics, IWA-IAHR Publishing, UK, 3:8, pp. 207–222.

[10] Hanson,J.; (1989), Order and structure in urban design: The plans for the rebuilding or London after the great fire, Ekistics.

[11] Hanson,J.; (1993), City of London space and role, Urban Design Group Quarterly, Issue 48.
[12] Hillier B, Penn A, Hanson J, Grajewski T, Xu J, (1993), "Natural movement: or, configuration and attraction in urban pedestrian movement" Environment and Planning B: Planning and Design 20, 29 - 66
[13] Hillier B. (1997), Space is the Machine: A Configurational Theory of Architecture, Cambridge University Press, Cambridge, UK
[14] Hillier B. and Hanson J. (1984), The Social Logic of Space, Cambridge University Press, Cambridge, UK.
[15] Hillier, B. 1996. Cities as movement economies, URBAN DESIGN International.1(1): 41-60
[16] Hillier, B., Penn, A., Hanson, J., Grajews ki, T., & Xu, J. 1993. Natural movement: Or, configuration and attraction in urban pedestrian movement. Environment and Planning B: Planning and Design, 20, 29-66
[17] Javadi A.A. and Rezania M. (2009), "Applications of artificial intelligence and data mining techniques in soil modeling", Geomechanics and Engineering, 1:1, pp. 53-74.
[18] Pepoinis J.; (1989), Space, culture and urban design in the modernism and after, Ekistics.
[19] Rezania M., Javadi A.A. and Giustolisi O. (2008), "An evolutionary-based data mining technique for assessment of civil engineering systems", Engineering Computations, 25:6, pp. 500-517.