

An Effective Approach for Land Use Management in the Palestinian territories

Azzedine Hani

Laboratoire Ressources en Eau et Développement Durable, Université Badji Mokhtar Annaba, 23000 (Algérie).
haniazzedine@hotmail.com

Abstract. The current understanding of the causes of land-use change is dominated by simplifications, which in turn, underline many environment-development policies that led to unsustainable situation in Gaza City. Therefore, this research aims at establishing an integrated land-use management framework for Gaza city based on cause-effect relationship within sustainability context. It also aims at establishing a prediction model for the relationship between the most influential socio-economic, environmental and institutional indicators that had shaped the current land-use status of the city for the period 1967-2003. The methodology presented in this research work, offers opportunities to simulate the future demand of the different land-uses based upon actual land-use conditions and other determinant factors. The determinant variables of land-use changes have been identified and prioritized using statistical analysis and Artificial Neural Network (ANN). The results were compared with other statistical techniques and expert opinions. ANN prediction model helped in drawing scenarios for future development. Combinations of socio-economic, environmental and institutional variables in addition to the actual land use for the last four years are used as a basis of land-use change explanations and modeling. These pathways indicate that land-use policies and projections for the future must not only capture the population indicators as the only drivers for land-use change but also account for the specific human resource development indicators and urban-environmental conditions. This recognition requires moving beyond some of the simplifications that persist in much of the current understanding of the causes of land-use change and its driving forces. The analysis of the local expert's opinions provide evidence support the conclusion that the simple answers found in population growth, poverty and infrastructure rarely provide an adequate explanations of land-use changes. Rather, social responses follow from changing economic conditions, mediated by institutional factors are the real causes for land-use changes in Gaza.

1. Introduction

Sustainable land-use management aims at integrating socio-economic, environmental and institutional factors into one framework that leads to formulate the appropriate policies and activities. Developing indicators framework is essential tool for tracking progress towards sustainable land-use. The selection of these indicators should be based on objective criteria and well-designed conceptual framework.

The researches done until now for developing a conceptual framework for modeling sustainable land-use management are still fragmented and incomplete attempts. Most of the reviewed researched focused on investigation the tools for special modeling of urban landuse rather than analyzing the interaction between factors that effecting the changes of urban land-use. Very few scholars described the relationship between the driving forces, environmental impacts and urban expansion in one framework [1]. Therefore, for the efficient analysis of land-use change problems, it is essential to establish a conceptual integrated framework that aims at achieving sustainable land-use management. It is essential to look for the appropriate modeling approaches and tools that could give acceptable results in analysis and predicting the available data. Traditional statistical models such as multiple regression analysis, principal component analysis and factor analysis have been very successful in interpreting socio-economic activities. However, traditional statistics are criticized as being ineffective in modeling spatial and temporal data. Attempts have also been made to develop spatially explicit models of land-use and land cover change by using several tools such as satellite imagery, aerial photography, GIS, remote sensing, agent-based modeling and ANN based modeling [2].

The research aims at achieving the following objectives:

- To establish integrated land-use management framework for Gaza city based on cause-effect relationship within sustainability context.
- To establish a prediction model for the relationship between the most influential socio-economic, environmental and institutional indicators, that had shaped the current status of the land-use in the city for the period 1967- 2003. This prediction model is the basis for drawing scenarios for future development.
- To develop recommendations for sustainable land-use management.

2. Material and Methods

Gaza City is situated in a strategic location between Asia and Africa, between the desert in the South and the Mediterranean climate in the North. The total area of the city is about 45,353 km², which is about 12% of the total area of the Gaza Strip. The city consists of 17 neighborhoods, varying in characters, which are formed by different residents. The environmental indicators had multi-dimensional implications and characterized by continuous deterioration. The increase of the amount of the generated wastewater, increased abstraction of ground water, increase the density inside the city and increase the amounts of the annual solid waste are adding more pressure on the limited natural resources.

The analysis plan followed in this research is designed to achieve two main objectives; firstly, selecting the most influential, socio-economic, environmental and institutional, indicators and secondly, modeling the relationship between these indicators and land-use changes. Several steps have been taken to achieve these objectives using different tools (Figure 1).

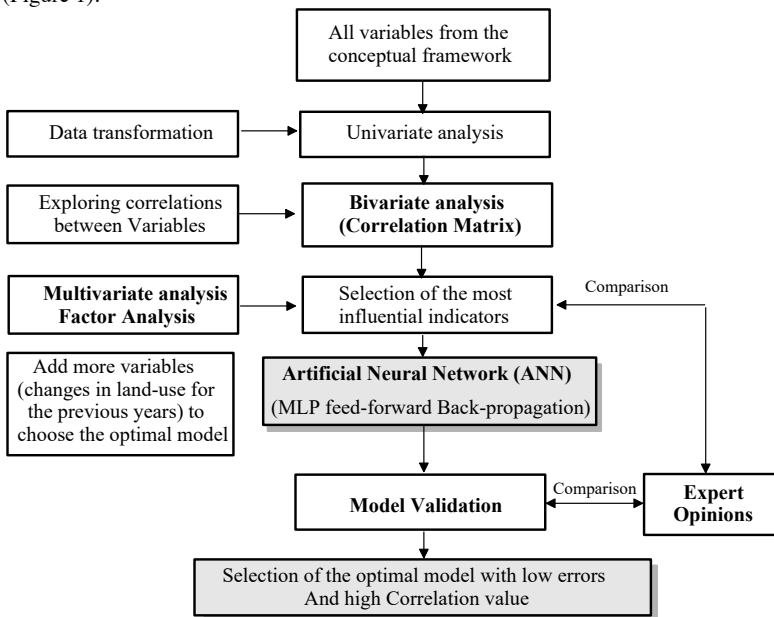


Fig. 1. Analysis Plan and Methodology.

3. Results

3.1. Statistical Analysis

Univariate, bivariate and multivariate techniques have been used to investigate how socioeconomic, environmental and institutional indicators are related to explain the land-use changes. Despite the fact that these statistical methods have several assumptions such as normal distribution of the data and linear relationships with the dependent variables, it was helpful in identifying the determinant indicators and factors that influence the changes in land-use of Gaza. The statistical analysis resulted with 10 indicators that have significant correlation with the different land-uses in Gaza and contributed by major role in shaping the land-uses for the last 37 years. These correlations vary from 0.7 to 0.95 in both directions (positive or negative correlations). These indicators are Log population under 19 years (P19), Log population density (PD), Log households connected to the wastewater (HWW), Log labour forces (LF), unemployment rate (UE), poverty rate (PR), households used cesspits for wastewater disposal (WWC), groundwater abstraction (GWA), groundwater salinity (TDS), and funds for development (FUND). Among these 10 indicators LF, PD and PR have strong correlation with all land-uses. Therefore these three indicators should be considered as the main driving forces behind land-use changes in Gaza.

The correlation coefficient (r) was calculated for each pair of variables to avoid using highly correlated variables (i.e. coefficient of correlation over 0.80) and to reduce the effect of multi-collinearity [3].

The correlation matrix reveals that population density has significant correlations among all the 10 determinate indicators. This confirms the fact that the densely areas are associated with more socio-economic and environmental

problems. Therefore population density is considered as the main driving force for land-use changes in Gaza. Also groundwater abstraction has correlated significantly with other indicators, while funds for development has a moderate to low correlations with other indicators.

Factor Analysis: the main applications of factor analytic techniques are to reduce the number of the independent variables and to detect structure in the relationships between these variables, that is to classify variables. Applying factor analysis to the data on Gaza city, with the 10 indicators resulted by using minimum eigenvalue method, it is noticed that only 2 factors were successfully extracted with eigenvalue more than 1.0 (Table 1).

Table 1. Eigenvalues using principle component extraction method.

Factor No.	Eigenvalue	% of the Total variances	Cumulative Eigenvalue	Cumulative %
1	6.98	69.82	6.98	69.82
2	2.25	22.48	9.23	92.30

From the second column (*Eigenvalue*) of the table above, we find the variance on the new factors that were successively extracted. In the third column, these values are expressed as a percent of the total variance. As we can see, factor 1 accounts for 69.8 percent of the variance which means that Factor one explained more than 72 percent of the actual changes in land-use of Gaza city, factor 2 accounts around 22.5 percent. The first two eigenvalues cumulated around 92 percent of the total variance, while the other 8 factors explained less than 8 percent of the changes in the land-use areas. Considering the first column of loading in the factor loadings matrix, shown in Table 1, indicates that this factor has a strong influence on the six indicators which are Log population density, Log households connected to wastewater network, Log households uses cesspits wastewater disposal, Log Groundwater abstraction, Log domestic water salinity and funds for development. It combined the social, environmental pressure and environmental impacts indicators. This factor is clearly identified the quality and quantity of the Gaza urban environment. Therefore it could be named as urban-environment management factor.

Considering the second column of factor loadings, it is clear that this factor has influence on four indicators and a moderate to weak influence on the remaining indicators. Since those four indicators focusing on human resource development indicators therefore this factor could be named as human resources development factor. This interpretation is consistent with the finding that this factor has a moderate effect on the other indicators due to the lower investment in the field of human resource development during the last 37 years. The slow improvement of the economical conditions in Gaza has insignificant impacts on development of social condition. These difficulties combined with the political instability in the area are hindering the efforts by the local institutions to minimize the pollution pressures nor they can provide remediation for the environmental negative impacts that threatening the public health.

Comparison with Expert Opinions and Judgments: the comparison of the statistical results with expert's opinions has significant value in evaluating the current understanding of local experts to the influential indicators that is driving land-use changes in Gaza. Therefore 23 local experts were identified as sources for individual indicator-ranking information. Experts were chosen according to their knowledge in the field of land-use management as well as their professional ability to provide ranking for the identified socio-economic, environmental and institutional indicators. Each expert was asked to rank these indicators (30 indicators) in order of its relative importance and relevance to the land-use management in Gaza City. The results of expert's opinions were accomplished through the circulation of the indicators-based questionnaire.

The results of the questionnaire were averaged within each indicator, and a preliminary ranking was computed. The results shows that the overall agreement between expert's and the statistical results was only 47%. The agreements were found in 14 indicators mainly population density, population under 19 years, unemployment rate, poverty rate and funds for development. In the other hand this analysis revealed that the environmental indicators were less estimated which resulted significant differences from the statistical analysis.

Therefore the statistical analysis compared to expert's opinions provide evidence support the conclusion that the simple answers found in population growth, poverty and infrastructure rarely provide an adequate understanding of land-use changes. Rather, individual and social responses follow from changing economic conditions, mediated by institutional factors.

Opportunities and constrains for new land-use are created by markets policies, increasing influence by regional factors. Various urban-environment conditions react to and reshape of the impacts of drivers differently, leading to specific pathways of land-use change. It is precisely these combinations that need to be conceptualized and used as the basis of land change explanations and modeling. This conclusion makes it more complicated to the traditional statistical models to produce an accurate prediction model for the changes of land-uses that could deal with complex relations between the indicators and its underlying factors. In addition, the collected data on the selected indicators showed non-linear relationships and non normal distribution.

Therefore, the next stage of modeling requires advanced techniques that provides a nonlinear function mapping of a set of input variables into the corresponding network output, without the requirement of having to specify the actual mathematical form of the relation between the input and output variables, it should has the versatility for modeling a wide range of complex non-linear phenomena.

3.2. Analysis and Modeling Using Artificial Neural Network (ANN)

ANN used here as analysis and predictor tool aims at exploring and modeling the relationship between the predictor variables (the determinant indicators, the inputs) and predicted variable (land-use indicators, the output), even if the relationship is very complex and not easy to articulate in the usual method of "correlations" or "differences between groups".

The advantages of the artificial neural network ANN approach are: its ability to handle nonlinear functions, to perform model-free function estimation, to learn from data relationships that are not otherwise known and, to generalize to unseen situations.

The output variables are fixed for all trained models, which are representing the percentages of land-uses, divided into five sub-indicators which are residential (RES), roads (ROD), industrial (IND), recreational (REC) and agricultural (AGR) indicators. The output data have been divided into three subsets; training (60%), testing (20%) and validation (20%). The results of these models, hereafter, called land-use models.

A total of 20 combinations of input variables presented in four groups are investigated.

The first group consists of nine models with a combination of the SEEI(t) variables and additional land use variables from the previous years $LU(t-1)$, $LU(t-2)$, ..., $LU(t-7)$.

The second group has four models with combination of SEEI variables from the previous 4 years {i.e. $SEEI(t)$, $SEEI(t-1)$, ..., $SEEI(t-4)$ }.

The third group has three models with combination of additional variables from land-use variables $LU(t-1)$, $LU(t-2)$, $LU(t-3)$ and $SEEI(t)$, $SEEI(t-1)$, $SEEI(t-2)$ and $SEEI(t-3)$ variables.

The fourth group consists of five models with combination of $SEEI(t)$, in addition to the changes in land use expressed by $\Delta LU(t)$. Additional $\Delta LU(t-1)$, ..., $\Delta LU(t-5)$.

The optimal ANN model's structure that resulted in minimum error and maximum efficiency during both training and testing was selected for validation. This experiment was done for developing MLPs; then, based on the performance of various trained MLPs with different input combinations, the best network were selected.

The data for each input combination were divided into a training set, a validation set, and a testing set. The training set consists of a set of examples used only for learning, (i.e. to fit the weights of the network). The validation set is a set used to adjust the network parameters such as network architecture (for example number of hidden layers and neurons, or number of training cycles). The test set is a set of examples used only to assess the generalization performance of a trained neural network, (external valuation). Hence, MLP was applied to the prediction of land-use areas in four phases: (1) design of the network and of inputs from historical data; (2) network training using a subset of inputs; (3) testing of the neural network using the full data set of the inputs; and (4) using the information from the neural network to forecast changes. To avoid over-training of the network, the neural network was trained with a different set of data by changing the trained set of data in random order. Therefore about 60% of the investigated data set was chosen as training data set (i.e. 21 cases), while other 20% (i.e. 8 cases) chosen as cross-validation data set and the rest cases 20% (i.e. 8 cases) were chosen as testing data set. This procedure has been repeated several times with different data inputs. In order to select the optimal model, about 16 models have been produced. Each model used different data sets have a significant influence on the predicted flow.

ANN Model used successfully as an optimization tool in a number of case studies. ANN well suited to the task of selecting an appropriate combination of inputs to a model as it has the ability to search through large numbers of combinations, where interdependencies between variables may exist. The main steps in the optimization process aims at determining the input variables, which have a significant influence on the predicted land-use changes. Consequently, the followed procedure analyzed different combinations of the influential socio-economic, environmental and institutional (SEEI) indicators in addition to the antecedent land-use values (i.e. $LU_{(t-1)}$, $LU_{(t-2)}$... $LU_{(t-n)}$) or the SEEI indicators for the previous years (i.e. $SEEI_{(t-1)}$, $SEEI_{(t-2)}$, ..., $SEEI_{(t-n)}$). Input variables including combination of the data from previous years on both LU and SEEI are also trained.

For the network structure identification, the current study employed a procedure of time series of land-use changes. This procedure suggests training a network with minimum input variables initially (i.e. the 10 SEEI indicators with 37 data sets), and systematically increasing the input vector dimension by adding more variables from the previous year and training the new network. This procedure, though adding complexity to the model, considers the ability of additional input variable to capture any nonlinear residual dependencies.

The number of hidden neurons in the network, which is responsible for capturing the dynamic and complex relationship between various input and output variables, could be identified by various trials [4]. The trial-and-error procedure started with one hidden neurons initially, and the number of hidden neurons was increased up to 4 during the trails. The small number of hidden neurons refers to the limited number of observations. For each set of hidden neurons, the network was trained in batch mode to minimize the average-square error at the output layer. In order to check any over-fitting during training, a cross-validation was performed by keeping track of the efficiency of the fitted model. The training was stopped when there was no significant improvement in the model's efficiency, and the model was then tested for its generalization properties.

The selection process of the desired optimal network model is composed of two consecutive stages, which indicate the values of efficiency. In the first stage, statistical accuracy measures such as coefficient of determination (R-square), Average Squared Error (ASE) and the Mean Absolute Relative Error (MARE) on both training and testing data sets to filter out the most promising optimal networks. The values of both ASE and MARE close to zero indicate a better performing model. The values of R^2 range from 0 to 1, with higher values close to 1 indicating better model performance. Statistically, an optimal network is defined as the one with the best overall accuracy measures. In the second stage, the predicted and actual outputs graphical responses for both training and testing sets for the selected most promising networks. Based on the overall graphical evaluation of each model's performance, the absolute optimal network can easily selected.

Group 1: The model performance continues to increase significantly with the additional antecedent land-use information to the SEEI(t) variables. Note that a combination of time-series data on land-uses and SEEI(t) variables in the input vector significantly improve the performance among the inputs models in the first group. The performance of model (6) indicates that addition of antecedent land-use information for more than 4 year is not significant in the optimal input vector. The best performance was achieved for Model 5.

This model has efficiency of 83% to 97% for the training data sets and 91 to 97% for validation data sets.

Group 2: Among the second group, the performances of the model 11 indicates that addition of antecedent SEEI information for more than 3 year did not significantly improve the performance of the model. Therefore, model 10 is best model among this group, however, this model has performance ranged between 71% to 83%, which is lower than the optimal model in group 1.

Group 3: Models 13, 14 and 15 in the third group presents the combination of additional variables from both land uses and SEEI indicators from the previous years. These additional inputs did not improve the model's performance significantly.

Group 4: Similar to the first group the model performance continues to improve for additional 4 years of changes in land use. Model (19) has better performance among the other models among this group, however, its results still lower than it found in the model number (5).

The graphical comparison of the different models' performance is given in Figure 2. It is obvious between the developed ANN model 5 has close values of R-square in both training and testing phases. This clearly indicates the powerful function approximation features of the back-propagation in the multi-layer perceptron ANN method.

The average squared error (ASE) and the mean absolute relative error (MARE) for all landuse categories have been calculated. The results confirmed the findings of the comparison between the observed and predicted percentages of land-use categories. It is obvious that the model 5 has the smallest values for both ASE and MARE.

The optimal network found was MLP with 30 input predictive variables and one hidden node. This resulted a model that has a good performance for land-uses with average performance ratio of 93% and ASE of 0.003. The input vector is represented by the 10 determinate factors and the land-use areas for the previous 4 years.

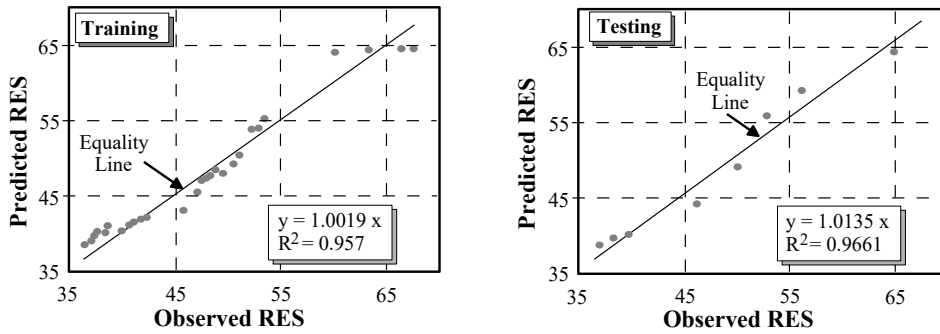


Fig. 2. Observed versus Predicted Residential Areas (Training and Testing).

The land-use percentages model can be represented by the following compact form:

$$LU(t) = ANN \left\{ \begin{array}{l} P19_{(t)}, PD_{(t)}, HWW_{(t)}, LF_{(t)}, UE_{(t)}, PR_{(t)}, WWC_{(t)}, GWA_{(t)}, TDS_{(t)}, FUND_{(t)}, \\ LU_{(t-1)}, LU_{(t-2)}, LU_{(t-3)}, LU_{(t-4)} \end{array} \right\}$$

Where $(LU)_t$ is the Land-use areas at time (t) which is equal to the sum of $(RES + ROD + IND + REC + AGR)$ at time (t) .

4. Conclusion

In this research, a sustainability approach for land-use management has been developed based on cause effect relationship. The sustainability approach gives insight into the management of land-use as a function of the complex interaction of a large number of socioeconomic, environmental and institutional factors. The effectiveness of such an approach depends on finding the non-traditional ways of dealing with land-use management. The methodology presented in this research work, offers opportunities to simulate the future demand of land-use changes based upon actual land-use conditions and other determinant factors.

The determinant variables of land-use changes have been identified and prioritized using multi-criteria analysis with Artificial Neural Network (ANN) compared with other statistical techniques and expert opinion. The main findings of these analysis techniques are:

- Population group under the age of 19 years old followed by population density, access to wastewater network, percentage of labour forces, unemployment rate and poverty rate are the most socio-economic driving forces for land-use change.
- Percentages of households use cesspits for wastewater discharge, and ground water abstraction are the main pressing environmental pollution sources that associated with land-use changes. The salinity of domestic water is the most representative variable of the environmental impacts of land-use changes.
- Policies, regulations and funds allocated for development are the most institutional measure that affects the land-use changes.

References

1. A. Veldkamp, P. Verburg, Elsevier Ltd. All. **72** 1–3 (2004)
2. M.S. Eila, *Thèse de Doct., de l'USTL, France*. 194 pp. (2005)
3. J.F. Mas, H. Puig, J.L. Palacio, A. Sosa Lopez, *Modeling Environmental Modeling & Software* **19**, 461–471 (2004)
4. H.R. Maier, and G.C. Dandy, *Environmental Modelling & Software* **15**: 101–124 (2000)