Journal of Coastal Research	SI 39	1515 - 1519	ICS 2004 (Proceedings)	Brazil	ISSN 0749-0208

Litter Categorization of Beaches in Wales, UK by Multi-layer Neural Networks

C. E. Balas†; A. T. Williams ‡; A. Ergin∞ and M.L. Koc §

†Civil Engineering Dept. Gazi University
Ankara, 06570, Turkey
cbalas@gazi.edu.tr

‡Built Environment, Swansea Institute, University of Wales, Swansea, UK
allan.williams@virgin.net

∞ Civil Engineering Dept.
Middle East Technical Univ.
Ankara, 06570, Turkey
ergin@metu.edu.tr

§ Civil Engineering Dept.
Cumhuriyet University
Sivas, 58140, Turkey
mkoc@cumhuriyet.edu.tr

ABSTRACT

BALAS, C. E.; WILLIAMS, A.T.; ERGIN, A. and KOC, M.L., 2006. Litter categorization of beaches in wales, UK by multi-layer neural networks. Journal of Coastal Research, SI 39 (Proceedings of the 8th International Coastal Symposium), 1515 - 1519. Itajaí, SC, Brazil, ISSN 0749-0208.

Litter categories and grades of Welsh beaches were satisfactorily predicted by multi-layered feed forward neural networks and fuzzy systems, which are artificial intelligence techniques. Neural network structures with hidden layers consisting of 40 neurons of uni-bipolar sigmoid functions were constructed for Welsh beaches and they were trained by supervised (conjugate gradient) learning algorithm to predict the number of litter items and categories from data obtained by 157 litter surveys carried out for 49 beaches in Wales, UK (including the most attractive tourist beaches of Tresaith, Aberporth, Port Eynon, Trecco Bay, Sandy Bay, Swansea Bay, Rest Bay, Lavernock, Goodwick, Amroth Castle, Rhyl Prom and Porthdafarch). The input data for trained neural networks were litter items in general litter category, and the network could predict items in remaining seven categories by learning the relation among them and considering main litter sources in UK (river, shipping, fishing, beach users and sewage related debris). These high-speed predictions saved on field efforts as fast and reliable estimations of litter categories were required for management studies of these beaches. Fuzzy systems were also used to incorporate additional information inherent in linguistic comments/judgments made during field studies and questionnaires distributed to beach users. The artificial intelligence model (ARIM) presented is a universal one to predict litter categories in different countries, which have various litter sources and beach user characteristics.

ADDITIONAL INDEX WORDS: *Artificial intelligence, fuzzy systems, beach grading*

INTRODUCTION

Litter on beaches on a worldwide basis is currently one of the main global problems. Clean up schemes have raised public awareness but has not stopped the tide of litter that strands the world's beaches. Marine litter has been defined as 'any manufactured or processed solid waste material (typically inert) that enters the marine environment from any source' (THORNTON and JACKSON, 1998). The effects of litter on beaches and to humans/animals have been widely and variously documented. For example, MORRISON and MUNRO (2000) on waste disposal; SPEAR *et al.* (1995) on the effect on seabirds; HALL (1998) on fishing; WILLIAMS *et al.* (2000) for hazardous materials; MARSHALL and ELLIOTT (1998) for medical waste; NELSON *et al.* (2000) on the effects on tourism. Examination of the literature regarding this subject presents a diverse mix of objectives. In scientific studies, these will have major influences on the methodology applied and this is certainly the case with marine litter surveys. As a result, a broad variety of techniques exist to describe and measure litter, which are not directly comparable. For example, litter can be categorised according by size (RIBIC, 1998), weight (FROST and CULLEN, 1997), number of black bin-bags collected or composition (DIXON and HAWKSLEY, 1980). There is as yet no single accepted methodology for assessing beach litter. Available methodologies of marine litter assessment (BALAS *et al.,* 2003) cannot predict the number of litter items and grades of beaches based on previous surveys of litter data and questionnaires. These methods require the actions of counting, classification, surveying and evaluation carried out at various beaches for all litter categories. Many of the beach grading systems in operation, such as the EA/NALG (2000) scheme, assign a classification to a beach after sampling a small part due to logistical, time and financial constraints. This methodology involved taking 100m stretch of beach, counting litter items according to 7 distinct categories and grading the beach (WILLIAMS *et al.,* 2003).

The artificial intelligence model developed satisfactorily predicted litter categories based on general litter category for Turkish (Antalya) beaches (BALAS *et al.,* 2003). Now the model was applied to Welsh beaches, and it is seen that, model predictions had high correlations, although the main litter sources in UK were determined as river, shipping, fishing, beach users and sewage related debris (TUDOR and WILLIAMS, 2001), but the main beach litter source in Turkey was "beach users", as obtained from principle component analysis. Therefore, this artificial intelligence model is a universal model that will save on field effort when fast and reliable estimations of litter categories are required for management purposes. The developed model, which would lead to better management schemes concerning beach health, can perform future predictions of litter items, categories and safety and it also has economic implications.

ARTIFICIAL INTELLIGENCE MODEL (ARIM)

The artificial intelligence model (ARIM) developed for marine litter prediction involves two sub-models: 1) Neural network 2) Fuzzy system. Both of the sub-models can be used for litter prediction and grading of beaches depending on the type of data availaBle. Neural networks can predict litter grades from new data sets by learning the main characteristics of environmental changes; therefore they are similar to a biological neural system. However, if the database is limited and contains qualitative information, the fuzzy sub-model can be used to assess litter grades, since this sub-model can utilize uncertainties inherent in human knowledge. In other words, the fuzzy model can handle human based information such as experience and judgment and can consider qualitative data described by language, such as questionnaires distributed to beach users. The disadvantage of fuzzy model is that, it does not have the capability of learning like the neural network model (BALAS and KOC, 2002).

Neural Network Sub-model (ARIM-1)

Neural networks are simplified models of the human brain. Connections between neurons are axons and dendrites,

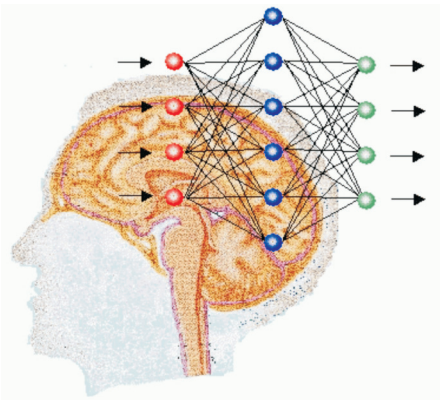

Figure 1. Artificial neural network analogy.

connection weights represent synapses, and the threshold approximates the activity in the soma as given in Figure 1.

The response of network: $y_j = f(\sum_{i=1}^{n} w_i x_i - t_j)$, where

y_j is the output of the jth neuron, x_i is the ith input signal, w_i is the connection weight from the ith input neuron to j^{th} neuron, t_j is the threshold or bias value of the jth neuron and f(x) is a nonlinear function. A multi-layer feed-forward neural network (MFF) has a layered structure as given in Figure 2. Hidden layers exist between the input (first) and the output (last) layer. Each layer has neurons that receive input from a layer below and send their output to units in a layer above. Activation of a hidden unit is a function of the weighted input and the threshold. A multi-layer feed-forward neural network is trained by a supervised learning algorithm with a set of chosen examples called a training data set and then tested by a data set. Neural networks are classified according to their activation phase as feed forward or recurrent, and according to learning phase as supervised or unsupervised.

In the Supervised learning algorithm, weight and bias factors were determined by minimizing the convergence criteria. The performance index (∇ P) is defined as the mean squared error:

$$\nabla P(\mathbf{w}) = \frac{1}{zN}\sum_{n=1}^{N} \nabla K_E(\mathbf{w}, n) \qquad (01)$$

The weight and bias updates are proportional to the performance index by total instantaneous squared error:

$$K(\mathbf{w}, n) = \frac{1}{2}\boldsymbol{\varepsilon}^T(n)\ \boldsymbol{\varepsilon}(n) \qquad (02)$$

where, N is the number of input and output vectors, n is the epoch number, is the output error at nth epoch, z is the number of neurons at the output layer, ∇ **K**(**w**,n) is the gradient vector of total instantaneous errors that have components associated with the weights of the hidden and output layers w^h and w^y, respectively for j=1 to s, i=1 to k and =1 to z, m=1 to s.

$$\nabla K(\mathbf{w}, n) = \left[\frac{\partial K}{\partial w_{11}^{h}} \cdots \frac{\partial K}{\partial w_{ji}^{h}} \frac{\partial K}{\partial w_{11}^{y}} \cdots \frac{\partial K}{\partial w_{\ell m}^{y}} \right] \qquad (03)$$

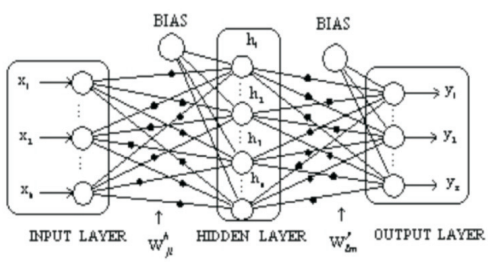

Figure 2. Multi-layer feed-forward neural network.

Fuzzy Sub-model (ARIM-II)

A fuzzy rule base characterizes the relationship between two or more variables in the general form of "**if**" antecedent proposition; "**then**" consequent proposition or a set of consequences that can be inferred**.** In the linguistic fuzzy system the antecedent and consequent propositions are fuzzy propositions. Fuzzy logic differs from binary logic in the way that the membership function in binary logic suddenly jumps from 0 to 1, while the membership function in fuzzy logic smoothly varies between the values of 0 and 1. Fuzzification is the transformation of input fuzzy sets into appropriate forms of fuzzy rule based system. In this system, logical relationship between the fuzzy input and output sets are revealed and quantified. The fuzzy rule base consists of conditional statements that describe the dependence of one or more linguistic variable on another. The analytical form of an if/then rule is the fuzzy relation called the implication relation, which can be defined over discrete universes of discourse:

$$U(x, y) = \sum_{(x,y)} g(x, y)/(x, y) \qquad (04)$$

where, is the membership function of the implication relation. Implication relations are obtained through different fuzzy implication operators, for example the Larsen implication operator is defined as:

$$g(x, y) = g_A(x) \cdot g_B(y) \qquad (05)$$

Fuzzy rules are connected by an aggregation operator of either "union" or "intersection" depending on the implication operator. Results obtained from the fuzzy rule based system are retransformed from internal fuzzy quantities (consequent) into numerical quantities or crisp outputs (y*) by the defuzzification methods (ZADEH, 1999). The Center of Gravity method is the most commonly used defuzzification method and is given as follows:

$$y^* = \frac{\sum y\ \mu(y)}{\sum \mu(y)} \qquad (06)$$

where, denotes the mean value.

MODEL APPLICATION TO WELSH BEACHES

Number of litter items in EA/NALG, (2000) categories (gross litter, sewage related debris-g (general), sewage related debris-cb (cotton buds), harmful litter-glass, harmful litter-other, accumulation and faeces) was predicted by using the number of litter items in general litter category for 49 beaches in the Principality of Wales, UK (including the most attractive tourist beaches of Tresaith, Aberporth, Port Eynon, Trecco Bay, Sandy Bay, Swansea Bay, Rest Bay, Lavernock, Goodwick, Amroth Castle, Rhyl Prom and Porthdafarch). Many of these beaches were subject to some form of cleaning regime, whether it was the whole beach or just certain strandlines. The accumulation of litter above the strandlines was greatly influenced by substrate, topography, vegetation, weather etc. Litter in this area provided useful information, especially with regard to long-term inputs, but is was not very indicative of the daily or new inputs. The area below the strandline on these beaches had been mostly shown to be almost free of litter and any time consuming surveys carried out below this line were generally ineffectual.

Number of litter items of 49 Welsh beaches was predicted by using multi-layer feed forward artificial neural networks trained by conjugate gradient learning algorithm using data obtained from 157 litter surveys. Litter amounts collected ranged from 0 to 1000-items/100 m stretch of beach. Litter items were graded from the best (Grade A) to worst case (Grade D) as shown in Table 1. Field measurements indicated that the most abundant litter item in terms of quantity was the general litter category. Sewage related debris, i.e. cotton buds was also in excess on these beaches (Table 2). The main sources of

Table 1. *Categories for grading a beach (EA/NALG, 2000).*

	Category	Type	A	B	C	D
1	Sewage Related Debris	General	0	1-5	6-14	>15
		Cotton Buds	0-9	10-49	50-99	>100
2	Gross Litter		0	1-5	6-14	>15
3	General Litter		0-49	50-499	500-999	>1000
4	Harmful Litter	Broken Glass	0	1-5	6-24	>25
5	Accumulation	Number	0	1-4	5-9	>10
6	Oil		Absent	Trace	Noticeable	Offensive
7	Faeces		0	1-5	6-24	>25

Table 2. *Range of litter items for 49 beaches in Wales, UK.*

Litter Categories	Maximum	Minimum
General litter	1000	0
Sewage related debris-G	46	0
Sewage related debris-CB	600	0
Gross Litter	181	1
Harmful litter-Glass	44	1
Harmful litter-Other	43	0
Accumulations	88	0
Faeces	13	0

riverine and beach debris in Wales, UK are sewage and ocean-based shipping related. Potentially harmful litter originates from trans-shipping. In rural areas the sewage system of the dispersed houses is directly connected to rivers without a treatment facility. Therefore cotton buds, condoms, feminine hygiene products can be frequently observed on beaches after a storm that washes away these items stranding on the side banks of rivers to beaches, which resemble foil hanging on a Christmas tree (BALAS *et al*., 2001). If there exists a certain relation between litter categories that can be identified by an artificial intelligence structure, the neural network can learn this relationship at its training stage.

For this case, amount of litter in most of the categories was in correlation with the general litter category, as items in general litter were increased, a similar trend was observed in other categories. For example, from Table 2, when there was no item in general litter, there was one (minimum value) counted in gross and harmful litter categories. When the number of items in general litter category was at its maximum value (1000), the number of litter items in 'sewage related debris' was counted as 600 (maximum value). This relationship between general litter and other categories was used for the construction of feed forward-multiple layer artificial neural networks for the prediction of litter items in other categories, as well as the beach grade. The number of litter items in general litter category was the input parameter to the neural network which then predicts the number of litter items in other categories given in Table 2.

At the training stage of artificial intelligence, new neural network structures with hidden layers consisting of 40 neurons of uni-bipolar sigmoid functions were constructed for Welsh beaches, since the network trained for Turkish beaches could result biased predictions due to different sources of litter categories observed in these countries. This new artificial intelligence structure was trained for Welsh litter database by using the conjugate gradient (CG) learning algorithm. This database contained more data (157 litter surveys) and additional subgroups for litter categories when compared to the Turkish one. The performance index for the training stage was selected as the minimum mean square errors of computations (BALAS and OZHAN, 2003). The minimum mean square errors of computations (MSE) at the end of iterations were calculated as 0.0672. Input/output values of neural networks were normalized between 0 and 1 by using the maximum and minimum values of the ranges in litter items.

At the testing stage, predicted number of litter items obtained from trained neural network is compared with measurements as given in Figure 3 where predictions had a high correlation value of R=0.979. Therefore, overall predictions of litter items in seven categories from general litter can be considered as acceptable. Here, as the performance criteria of neural networks, non-parametric Spearman rank correlation coefficients were utilized. The main sources of litter on these beaches were sewage related debris, river, shipping and fishing sources (WILLIAMS *et al.*, 2000). The artificial intelligence model (ARIM) learned the relation between litter categories and satisfactorily predicted litter items in remaining 7 categories using general litter data available for a specific Welsh beach. The beach litter sources in Wales had a more complex effect on beach pollution than in Turkey in which the main source of litter was beach users, and cigarette butts were the specific liter item that was in excess on Turkish beaches. Litter surveys in Turkey formed an independent cluster in principal component analysis with "roadside surveys" in Wales, UK, which were studies conducted mostly in rural areas (Figure 4). In this figure, Turkish beaches are denoted as S47 to S53, and roadside surveys in Wales, UK are designated by S46, S57 and S65. Other numbers show the survey sites in Wales. The principle component analysis transforms the original set of variables to the set of uncorrelated linear components in decreasing order of importance to obtain independent groups by clustering. As a result, the artificial intelligence sub-model presented has a universal application field to save on field efforts when fast and reliable estimations of litter categories are required for management purposes of beaches.

The fuzzy module of ARIM has the advantage to utilize uncertainties inherent in human knowledge, such as questionnaires applied to beach users. In the fuzzification process of the system inputs, which were the number of general litter and sewage related debris items, the grading criteria of EA/NALG (2000) was utilized. The fuzzy input sets and membership functions of these variables were obtained for the beach grading (A: excellent, B: good, C: average and D: worst) depending on the number of litter items measured at Welsh beaches. Fuzzy input membership functions of grading system for the category of sewage related debris and general litter were developed for this case study as given in Figures 5 and 6. The Cartesian product of fuzzy input sets established the fuzzy rule based system for Welsh beaches. Fuzzy output sets for the grading of litter categories were coded as 1,2, 3 and 4, respectively, in order to convert the grading system into numerical values as illustrated in Figure 7.

The fuzzy sets of grades given in Figure 5 and 6 were modified to assess the changes inherent in other litter categories and sources for Welsh beaches by decreasing a half grade in the input fuzzy subset of rules considering the maximum number of litter items, greater than a certain limiting value in related

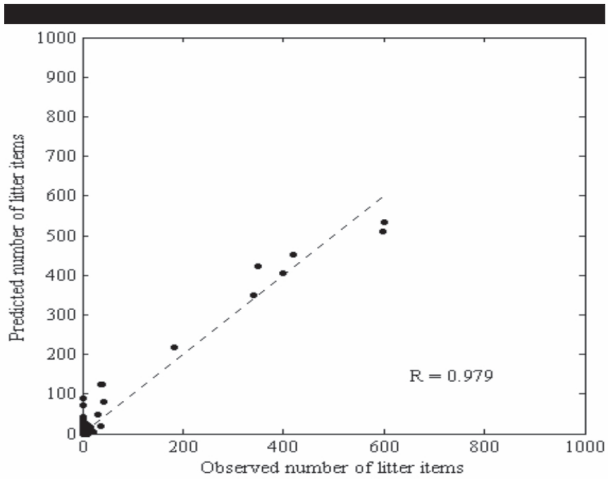

Figure 2. Multi-layer feed-forward neural network.

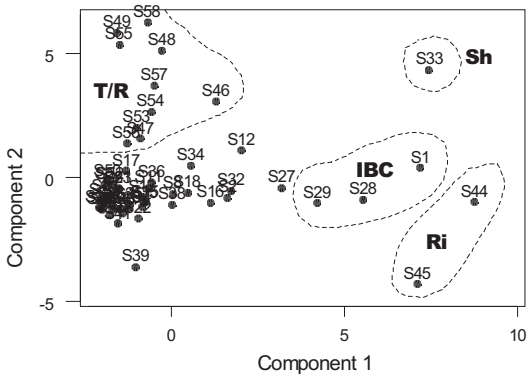

Figure 4. Comparison of principal components of beach survey sites in UK and Turkey (T/R= Turkish beaches/ roadside survey in UK; IBC= Inner Bristol Channel; Ri= River source litter; Sh= Shipping.

categories. In other words, fuzzy grades of general litter are decreased a half grade, when potentially harmful litter, gross litter and accumulations of litter exceed their limits. Similarly, the occurrence of faeces of non-human origin, which is very common in UK, affects the rules in the input fuzzy set for sewage related debris. At the testing stage of fuzzy sub-model, litter measurements were compared with predictions of fuzzy system, as shown in Figure 8. The fuzzy sub-model of ARIM satisfactorily predicted the grading of beaches, since the correlation coefficient between predictions and measurements was relatively high (R= 0.870). Predictions were performed for an average CPU (Central Processing Unit) time of 42 seconds within a standard mean error of ε =1% by using a Pentium III processor.

DISCUSSIONS

The overall predictions of the number of litter items in seven categories from general litter items at Welsh beaches can be considered as satisfactory when using both neural network and fuzzy sub-models of ARIM. The fuzzy system had a lower correlation between predictions and measurements, since this sub-model could not learn the relation between sources and categories, but it could assess the changes inherent in other litter categories and sources. In addition the fuzzy system could consider qualitative data described by language, such as questionnaires distributed to beach users. This sub-model could incorporate specific issues related to beach characteristics, litter assessment and oil pollution, which cannot be included in standard procedures.

The artificial neural network sub-model of ARIM could learn the relation between litter categories and satisfactorily predicted litter items in remaining 7 categories using general litter data available for a specific Welsh beach. The beach litter sources in Wales had a more complex effect on beach pollution.

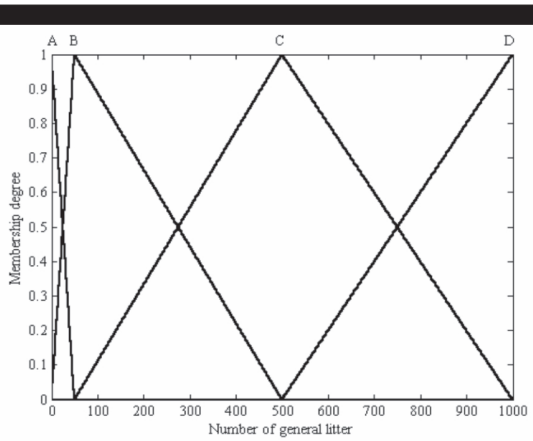

Figure 5. Fuzzy input membership functions of grading system for the category of general litter at Welsh beaches.

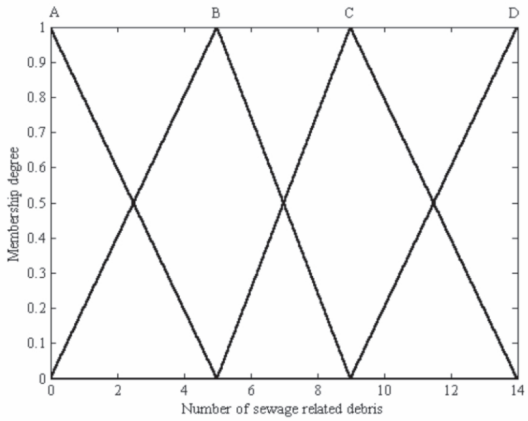

Figure 6. Fuzzy input membership functions of grading system for the category of sewage related debris at Welsh beaches.

CONCLUSIONS

Litter categories and grades of Welsh beaches were satisfactorily predicted by using two independent sub-models of the artificial intelligence model (ARIM) developed: 1) neural network 2) fuzzy system. Both of the sub-models can be used for litter prediction and grading of beaches as universal models depending on the type of data available. Neural networks can predict beach grades by learning the relation between litter categories and sources. Fuzzy sub-model can utilize uncertainties inherent in human knowledge to assess grading information for a database that contains qualitative information. For the neural network sub-model, a new network with hidden layers was constructed and trained by conjugate gradient learning algorithm from data obtained by comprehensive field campaigns carried out for 49 beaches in Wales, UK.

The input data for trained neural networks were litter items in general litter category, and the network could predict items in remaining seven categories by learning the relation among them and considering main litter sources in UK (river, shipping, fishing, beach users and sewage related debris). The beach litter sources in Wales had a more complex effect on beach pollution than in Turkey in which the main source of litter was beach users, and cigarette butts were the specific liter item that was in excess. Litter surveys in Turkey formed an independent cluster in principal component analysis with "roadside surveys" in Wales, UK, which were studies conducted mostly in rural areas. It is concluded that, the artificial intelligence model (ARIM) presented is a universal model to predict litter categories and grades in different countries, which have various litter sources and beach user characteristics.

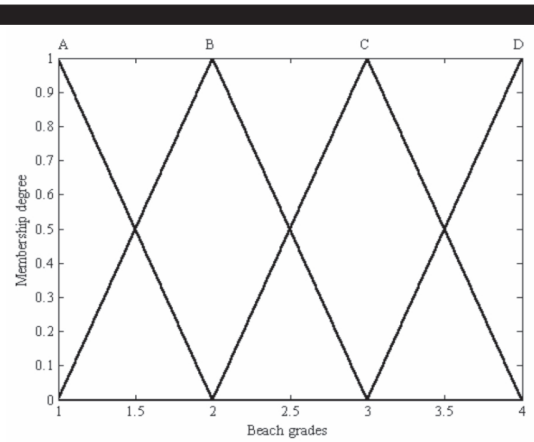

Figure 7. Fuzzy output membership functions for the grading of litter categories for Welsh beaches.

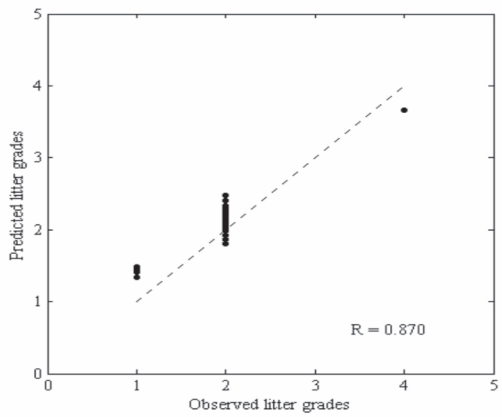

Figure 8. Comparison of beach grades of the field campaign with the predicted grades of fuzzy system.

LITERATURE CITED

BALAS, C. E.; ERGIN, A.; WILLIAMS, A.T. and KOC, L., 2003. Marine Litter Prediction by Artificial Intelligence, Marine Pollution Bulletin (in press).

BALAS, C. E.; WILLIAMS, A. T.; SIMMONS, S.L and ERGIN, A., 2001. A Statistical Riverine Litter Propagation Model, *Marine Pollution Bulletin,* 42 (11), 1169-1176.

BALAS, C.E. and KOC, L., 2002. Risk Assessment of Vertical Breakwaters - a Case Study in Turkey, *China Ocean Engineering*, 16 (1): 123-134.

BALAS, L. and OZHAN E., 2003. A Baroclinic Three Dimensional Numerical Model Applied To Coastal Lagoons, *Lecture Notes in Computer Science*, 2658: 205-212.

DIXON, T. and HAWKSLEY, C., 1980. *Litter on the beaches of the British Isles.* Report of the First National Shoreline Litter Survey. Marine Litter Research Programme, Stage 3, The Tidy Britain Group. 70pp.

EA/NALG, 2000. Assessment of Aesthetic Quality of Coastal and Bathing Beaches, *Monitoring Protocol and Classification Scheme,* Environmental Agency, UK.

FROST, A. and CULLEN, M., 1997. Marine debris on northern New South Wales beaches (Australia): Sources and the role of beach usage, *Marine Pollution Bulletin,* 34 (5): 348-352.

HALL, K., 1998. *Economic and Social Impacts of Marine Debris and Oil on Coastal Communities. Stage 1 Report.* KIMO and Napier University, UK.

MARSHALL, S. and ELLIOTT, M., 1998. Environmental influences on the fish assemblage of the Humber estuary, UK, *Estuarine Coastal And Shelf Science,* 46 (2): 175-184.

MORRISON, R.J. and MUNRO, A.J., 2000. Waste management in the small island developing states of the South Pacific: an overview. *Australian Journal of Environmental Management.* 6, 232-246.

NELSON, C.; BOTTERILL, D. and WILLIAMS, A.T., 2000. The beach as a leisure resource: measuring beach user perception of beach debris pollution. *Journal of World Leisure and Recreation.* Vol. 42 (1), 38-43.

RIBIC, C.A., 1998. Use of indicator items to monitor marine debris on a New Jersey beach from 1991 to 1996, Marine Pollution Bulletin, 36 (11): 887-891.

SPEAR, L. B.; AINLEY, D.G. and RIBIC, C.A., 1995. Incidence of Plastic in Seabirds from the Tropical Pacific, 1984-91: Relation with Distribution of Species, Sex, Age, Season, Year and Body Weight, ***Marine Environmental Research,*** 40, 2, 123-146.

THORNTON, L. and JACKSON, NL., 1998. Spatial and temporal variations in debris accumulation and composition on an estuarine shoreline, Cliffwood beach, New Jersey, USA, *Marine Pollution Bulletin,* 36 (9): 705-711.

TUDOR, D. T. and WILLIAMS, A. T., 2001. Some Thresold Levels in Beach Litter Measurement. *Shore and Beach,* 69 (4), 13-18.

WILLIAMS, A. T.; POND, K. and PHILLIPP, R., 2000. Chapter 12: *Aesthetic Aspects* (In), Monitoring Bathing Waters, (eds) J Bartrum and G Rees E, FN Spon, 283-311.

WILLIAMS, A. T.; TUDOR, D.T. and RANDERSON, P., 2003. Beach litter sourcing in the Bristol Channel and Wales, UK. *Water Air Soil Pollution,* 143 (1-4): 387-408.

ZADEH, L. A., 1999. Fuzzy Logic and the Calculi of Fuzzy Rules, Fuzzy Graphs, and Fuzzy Probabilities. *Computer and Mathematics with Applications*, 37, 35.