JNTS

JOURNAL OF NATURAL
AND TECHNICAL SCIENCES

STATISTICAL PROPERTIES OF ALBANIAN MOBILE PHONE COMMUNICATION NETWORKS

Eva NOKA and Fatmir HOXHA
University of Tirana, Faculty of Natural Sciences

ABSTRACT

Mobiles are a communication facility connecting people from around the world. As the use of mobiles has increased and the technology advanced, analyzing mobile phone communication networks is very interesting. The data collected from the mobile companies and the technological development made in the area help understand human interactions pattern and the society. In the present paper, we used data records of calls and SMS from a mobile phone company in Albania to construct and analyze these two communication networks: i) the directed network (DN), by linking callers or SMS senders and call or SMS receivers with directed edges and, ii) the undirected mutual network (MN) considering only the pairs of users with reciprocal communication events. In both cases, we assigned a weight on each edge, based on the number of communications occurred for the corresponding pair of users during the observed time period. Based on a carefully chosen set of topological and weighted measures, we present a detailed analysis of the networks under consideration, by exploring the component size, degree, strength, and weight distributions, topological and weighted degree assortativity and clustering coefficient. The correlations between these quantities were inspected, providing a better insight into the local structure and interactions, and the overall topology of the networks. This is the first study that attempts to analyze and discover the topological and weighted properties of mobile phone communication networks with Albanian users.
Keywords: mobile phone communication, network analysis, weighted networks

1. INTRODUCTION

Online social networks (Facebook, Twitter), E-mail, and mobile phones have transformed the way people communicate and interact with each other, providing connections with everyone all the time. The digital traces of social interactions, their frequencies and volume of communications generated by millions of people via all these communication channels offer a comprehensive picture of human habits. The management of such huge databases, network theory and statistical

tools have radically changed the scientific approach towards the understanding of our lives, the dynamic and performance of individuals, groups, organizations and societies (Watts, 2007; Lazer *et al*., 2009; Giles, 2012).

An anonymized data set of calls and text messages records generated by the users of a cellular operator in Albania is here used to explore the Albanian mobile phone communication networks. We performed a detailed analysis of the networks under consideration, based on a broad set of measurements that researchers have proposed to express and characterize the topological and dynamical features of the real-world networks (Newman, 2003; Barrat *et al*., 2004; Onnela *et al*., 2005; Boccaletti *et al*., 2006; Saramäki *et al.,* 2007).

2. MATERIALS AND METHODS

Data set The data consists of all call and text message (SMS) records among the subscribers of a cellular company which operates in Albania, during two nonconsecutive months or 61 days. This is a triple record $((i, j), t)$, i.e., the user i (*anonymized*) calling or sending a SMS to the user j (*anonymized*) at date/ time t. There are 4,323,433 records, where 3,910,079 are voice calls and 413,354 are SMS.

Construction of networks We constructed the original directed network denoted as the DN. The nodes represent the active mobile phone users, and there is a directed edge from a node i to a node j if the user i has called or sent a SMS to the user j at least once during the observed period. We excluded the loops from the network. There are 102,872 nodes and 243,197 edges in DN.

An undirected mutual network (MN) is obtained from the DN network as follows: each pair of edges (i, j) and (j, i) is substituted with an edge ij in the network MN, and all other edges are eliminated. The nodes with degree zero are removed. By this restriction, we only retain reciprocated ties and disregard the pairs of nodes connected by unidirectional edges, which mostly correspond to single communication events, suggesting that these pairs of users might not know personally each-other or the callers or SMS senders are more likely to be telemarketers or costumer service lines etc. Therefore, the mutual network MN should reflect a more realistic network of social interactions between the involved individuals. This filtering leads to the removal of about 44% of the total edges and 48% of the total nodes in the DN, remaining 53,453 nodes and 68,055 edges in the MN.

We assign a weight on each edge of both networks DN and MN to quantify the strength of a connection. In the DN, the weight w_{ij} of a directed edge (i, j) is the total number of calls and text messages initiated from the user i to the destination, the user j over the observed time period, whereas in the MN the weight w_{ij} of the edge ij is the total number of communication events (calls or

SMS) occurred between the users i and j.

Fitting distributions All the plots are in log-log scales unless otherwise specified. The scaling exponent of the power law distribution that best fit the component sizes is found using the Maximum Likelihood Estimation (MLE) as given in (Clauset *et al*., 2009). The best fitting parameters of the DPLN distribution are found with the MLE method, using the fitdistrplus package in R referring to the definitions and the estimation method proposed by Reed and Jorgensen (2004) and further information is provided in Noka (2014).

The exponent of the power law If there is a power law dependence of some data y on x of the form: $y = cx^{\alpha}$, the plot of the function y against x in doubly-logarithmic scales is a straight line with slope α. In the cases we treat here, we do as follows: We bucketize the x axes in logarithmic binning (Newman, 2006). For each bin we calculate the mean of y values, and plot them versus the midpoints of the bins in log-log scales. Using a linear fit, we find the exponent of the power law, if there is such dependence. All of the exponents are fitted with $R^2 > 0.95$.

3. RESULTS

Connected components

Both constructed networks are fragmented in a large number of connected components where the largest one contains the most of nodes as a giant component. In the directed network DN, 91,586 or 89% of the total nodes belong to its giant component, denoted as the gcDN. The giant component gcMN of the mutual network MN has 39,639 nodes or 74.2% of the total nodes. The mutual network MN is more fragmented than the DN, due to the removal of unidirectional edges. Table 1 summarizes the aforementioned information giving also the number of connected components and the number of nodes in the second largest component N_{SLC} for both networks.

Table 1. Sizes of the networks and their giant components

Network	Nodes	Edges	Connected components	The giant component		N_{SLC}
				Nodes	Edges	
DN	102872	243197	4307	91586	233970	21
MN	53453	68055	5002	39639	58718	26

The component size distribution $P(S)$ is the probability that a randomly chosen component in a network has S nodes. Figure 1A depicts the empirical component

size distributions plot for the networks DN and MN excluding their giant components. Both distributions exhibit power law distributions of the form $P(S) \sim S^{-\alpha}$, with exponents α_{DN}=3.52, α_{MN}=3.298.
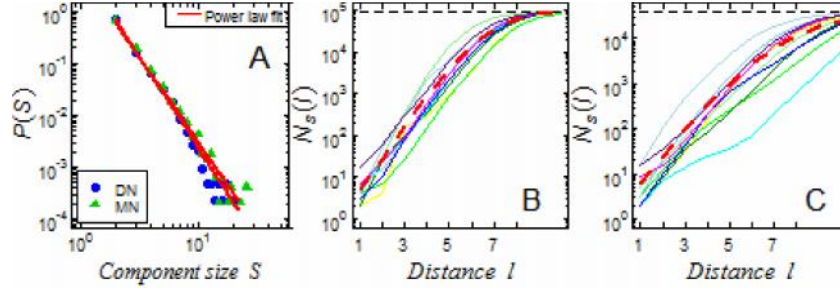

Fig. 1. Component size distributions (A). Number of nodes $N_s(l)$ for several samples and their average versus the distance l in the gcDN (B) and gcMN (C).

Snowball sampling is a method for obtaining samples of the large networks (Lee *et. al*., 2009). It starts from a randomly chosen node *v*, and includes all the nodes within a distance *l* (length of the shortest path) from the source node *v*. For a randomly selected set of source nodes, we obtain some samples using this method in each of the giant components gcDN and gcMN. By denoting as $N_s(l)$ the number of nodes in a sample, we plot each of them (solid lines) and their average (dashed red line) as a function of the distance *l*=1,2,3,... in log-linear scales. The illustrated results are shown in the panels B and C of Figure 1 for the gcDN and gcMN, respectively. The dashed red lines show an exponential growth up to the distance *l*=6, followed by a slower increase rate to the total number of nodes of the respective networks (dashed black lines), indicating the presence of the small-world effect (Watts and Strogatz, 1998) in both giant components gcDN and gcMN.
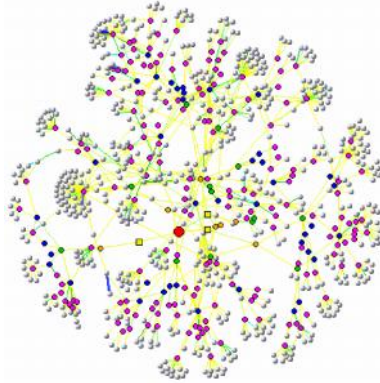

Fig. 2. A sample of the network gcMN, for *l* = 6.

A sample of the giant component gcMN, for l=6 is in figure 2 depicted. The color of nodes (yellow-green-blue-magenta-orange-lavender) corresponds to their distance from the source node colored in red. The lavender nodes dominate over all other nodes, whereas the neighbors of the red node are only 3 (yellow nodes). The width and the color (from yellow-green-cyan-blue to red) of the edges are drawn based on their weights.

Basic properties of the nodes and edges. The basic characteristic of the nodes in a network is the degree distribution denoted as $P(k)$, which gives the probability that a randomly chosen node has degree k. Panels A and B in figure 3 show the degree distributions of the networks DN, MN and those of their giant components gcDN, gcMN, respectively. The degree distributions of directed networks DN, gcDN have a fatter tail than those of the mutual networks MN and gcMN. The maximum degree of the gcDN is k_{max} = 1010, and the gcMN has k_{max} = 77, which is a more "reasonable" number of contacts for a single individual. The average degree of the gcDN and gcMN are $\langle k \rangle$ = 5.1 and $\langle k \rangle$ = 2.96 respectively, meaning that a user averagely communicates with about 5 persons, while he communicates reciprocally with about 3 of them.

All these distributions can be well fitted by the Double Pareto Lognormal Distribution DPLN (α, β, v, τ) with probability density function:

$$f(x) = \frac{\alpha\beta}{\alpha + \beta} \left[x^{-\alpha-1} \exp\left(\alpha v + \frac{\alpha^2\tau^2}{2} \right) \Phi\left(\frac{lnx - v - \alpha\tau^2}{\tau} \right) + \right.$$

$$\left. x^{\beta-1} \exp\left(-\beta v + \frac{\beta^2\tau^2}{2} \right) \Phi^c\left(\frac{lnx - v + \beta\tau^2}{\tau} \right) \right],$$

$$x > 0, \alpha, \beta, v, \tau > 0.$$

Φ is the cdf of a standard normal distribution and $\Phi^c = 1 - \Phi$ its ccdf. The estimated best fitting parameters are as follows:

DN $\alpha = 8.28, \beta = 0.1, v = 2.1, \tau = 0.76$ and for its giant component gcDN $\alpha = 8.3, \beta = 0.1, v = 2.2, \tau = 0.75$.

MN $\alpha = 9.4, \beta = 0.1, v = 1.4, \tau = 0.65$, and for the giant component gcMN $\alpha = 9.17, \beta = 0.2, v = 1.5, \tau = 0.65$.

Each network and its corresponding giant component present approximately identical degree distributions. From now on, we will focus on the analysis of the giant components gcDN and gcMN.
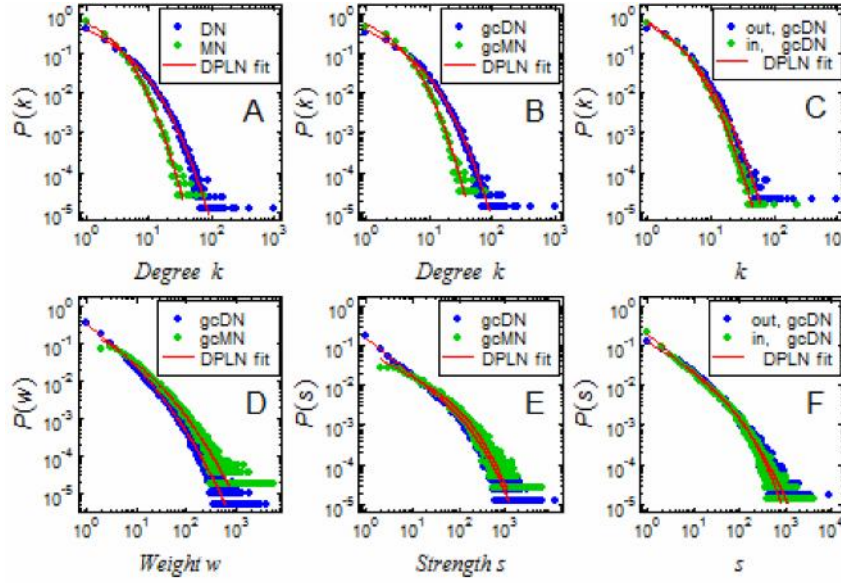
Fig. 3: Degree, strength, weight distributions fitted with DPLN distribution.

The panel C of figure 3 shows the in-degree and out-degree distributions of the gcDN fitted by DPLN distributions with the corresponding parameters: $\alpha^{in} = 8.66, \beta^{in} = 0.2, \nu^{in} = 1.3, \tau^{in} = 0.75$,
$\alpha^{out} = 7.9, \beta^{out} = 0.2, \nu^{out} = 1.6, \tau^{out} = 0.79$. It is observed that the nodes with very high degrees in the gcDN are mostly the contribution of their out-degree indicating that these hubs might be telemarketers or customer service numbers.

The communication intensity of any pair of users reflects the strength of their relationship, and numerically is characterized by the weight of the corresponding edge in the networks under consideration. Not only the node degrees, but also the tie weights show a heavy-tailed distribution in both networks gcDN and gcMN, as seen in Figure 3D. This indicates a strong heterogeneity in the way people choose to distribute their communications across their social contacts. Most of edges have small weights (weak ties) and only a few of them have high weights (strong ties) (Granovetter, 1983; Onnela *et. al.*, 2007a). Two connected users have reciprocally exchanged averagely 44 calls or SMS and maximally 5572 calls or SMS. The DPLN distribution fits well also the weight distributions in both networks gcDN and gcMN with the best fitting parameters as follows: for the gcDN $\alpha = 4.1, \beta = 0.001, \nu = 2.7, \tau = 1.42$, and for the gcMN $\alpha = 4.1, \beta = 0.3, \nu = 3.28, \tau = 1.47$.

The strength of a node i, $s_i = \sum_{j \in N(i)} w_{ij}$ is an extended definition of the node degree. In the networks under consideration, the strength of a node is the total number of calls or text messages made or received by a certain user during the observed time period. This quantity measures the volume of communications that a user performs and his importance (Barrat $et.$ $al.$, 2004). Strength distributions for the networks gcDN and gcMN are plotted in Figure 3E. For the directed network gcDN, we distinguish the incoming and outgoing strength of nodes by summing up the weights of incoming and outgoing edges respectively, and their distributions are plotted in Figure 3F. Obviously, there is a large variability of the volume of communications that people perform through their mobile phones in each direction too. Although the node strength is averagely 81.3 (1.3 events per day) in the gcDN and 130.6 (2.1 events per day) in the gcMN, the maximum value is 12402 (203.3 events per day) and 6651 (109 events per day) respectively. All the strength distributions can be fitted very well with DPLN distribution with the following parameters: for the gcMN $\alpha = 7.56, \beta = 0.3, \nu = 5.3, \tau = 1.05$,   for   the   gcDN, $\alpha = 7.2, \beta = 0.05, \nu = 5.5, \tau = 0.99$05, for the incoming strength in the gcDN, $\alpha = 7.32, \beta = 0.01, \nu = 4.97, \tau = 0.9$, and for the outgoing strength $\alpha = 7.1, \beta = 0.35, \nu = 4.1, \tau = 1.45$.

Similar pictures of the degree, strength and weight distributions observed in both networks gcDN and gcMN suggest the existence of mutual dependences between these quantities, leading us to a more detailed analysis below.

Degree, strength and edge weight correlations Degree correlation is an important parameter to better understand the topology of the networks. We want to investigate if there is a relation between the number of incoming and outgoing neighbors for a single node in the directed network gcDN. The scatter plots of in-degrees k-^{in} and out-degrees k-^{out} per node are in Figure 4A depicted. The blue dots are the medians by logarithmic binning (see Materials and Methods). It is observed the power law dependence $k^{out} \sim (k^{in})^{\alpha}$ with α=1.045. This means that the more popular individuals (those receiving many calls or text messages) tend to contact more people.

A measure of degree-degree correlation is *the average nearest neighbors degree* of a node i:

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in N(i)} k_j$$

where $N(i)$ is the neighborhood of node i. By calculating the average of the $k_{nn,i}$ values over the nodes with degree k, one can get the average degree of
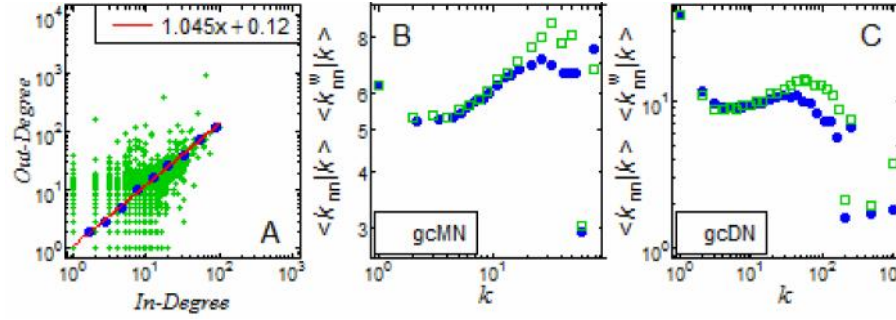
Fig. 4. In-Out degree correlation in the gcDN (A). Degree and weighted degree correlation for the gcMN (B) and the gcMN (C).

the nearest neighbors denoted as $\langle k_{nn} | k \rangle$. This quantity can be also expressed as $\langle k_{nn} | k \rangle = \sum_{k'} k' P(k' | k)$, where $P(k'|k)$ is the conditional probability that a node with degree k' is connected to a node with degree k. In absence of degree correlations, $P(k'|k)$ does not depend on k, and neither does , hence (Pastor-Satorras *et al*. 2001). On the contrary, if increases as a function of k, high-degree nodes have a lager probability to be connected to other high-degree nodes. In this case, the network is said to show *degree assortative mixing* (Newman, 2002). Barrat et al. (2004) has introduced *the weighted average nearest neighbors degree* of a node *i* in the weighted networks: . By averaging the values over the nodes with degree k, the behavior of the function marks the weighted assortative or disassortative properties, considering the interactions among system's elements. In Figure 4B and C, we plot and as functions of degree k for the networks gcMN and gcDN respectively. For the gcDN, we do not consider the direction of edges. In the mutual network gcMN, both functions and have an increasing trend with k up to the $k \approx 27$ and $k \approx 34$, respectively. The nodes with degree larger than $k = 27$ constitute only 0.08% of the total nodes. Therefore, the mutual network gcMN shows the degree (weighted) assortative properties, indicating that the users preferentially establish their relationships with individuals similar to themselves. This human tendency, known also as *affinity* or *homophily* emerges also with respect to a variety of attributes such as health or habits (Christakis and Fowler, 2007), tastes and interests (Lewis *et al*., 2008), age or race (Mollica *et al*., 2003). In the directed network gcDN, the function shows an increasing trend along a wider range of the k values than the , but the nodes with degree within this range constitute only 18.3% of the total nodes of the network. Thus, gcDN exhibits mostly disassortative mixing pattern.

We can also study the correlations between nodes strength. The average

nearest neighbors strength of a node i is defined as . By averaging values over the nodes with strength s, the average strength of nearest neighbors is obtained. We plot as a function of the strength s in Figure 5A for the networks gcDN and gcMN. Both curves present qualitatively similar pictures. The function shows an independence for small s values, followed by a slightly increase with the growth of s and a more obvious growth from $s \approx 800$ onwards. The nodes with strength greater than 800 establish only 2.5% of the total nodes in the gcMN, and 1.1% in the gcDN. Furthermore, the strength of such nodes is dominated by the weight of a single edge ij, thus, $s_{ij} \approx w_{ij} \approx s_{ji}$. These results suggest that in both networks, the strengths of two adjacent nodes are mostly uncorrelated.

The similarity of the strength distribution $P(s)$ with the degree distribution $P(k)$ in both networks, as seen above, leads us to analyze the relations between the node strength and the node degree in the networks gcDN and gcMN. If the weights of the edges incident to a node would not correlate with its degree, then with $\alpha = 1$, because , where is the average weight of the network. By plotting the average node strength as a function of degree shown in Figure 5B, we find that the node strength follow a power law growth up to $k \approx 45$ in the gcMN, with , and for larger k values the strength saturates. In the gcDN, the power law dependence exists with , up to $k \approx 25$, and for larger k values we observe a clear deviation from the linear approximation. These observations indicate the existence of strength-degree correlations and suggest that the individuals with many contacts averagely dedicate less calls or SMS per contact than those with few connections. The node strength conditional on in/out degree for the gcDN presents similar results (Figure 5C; the green dots are shifted vertically by a factor of 10 for better visibility). For small k values, both curves exhibit power laws, where $\alpha^{in} = 1.15$ and $\alpha^{out} = 0.94$, and deviations are observed for greater values of k.

In addition to the strength-degree correlations, we plot the average strength product conditional on degree product , shown in Figure 5D for both networks gcDN, gcMN. In the case of absence of correlations, we would expect . Then with $s = 1$. On the contrary, it is found that and .

Similarly, we analyze the average weight conditional on degree and strength product, and respectively. The corresponding results are presented in Figure 5E and F. In the mutual network gcMN, the exponent of the power law of the form , is , which means that the strength of a tie (edge weight) between two individuals does not depend on the absolute number of friends that each of them has. Despite the slight increase of observed for small values , practically the network gcDN also shows an independence of the edge weight versus the degree product. In contrast, the rapidly increasing curves in Figure 5F indicate the presence of correlations between the edge weight and the strength product
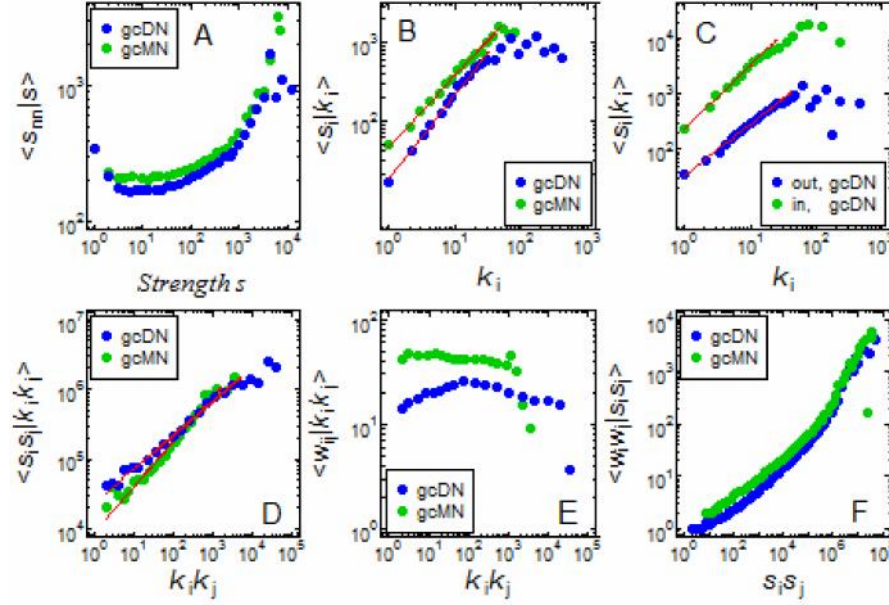
Fig. 5: Average nearest neighbor strength (A). Average strength conditional on degree (B). Average out/in strength conditional on degree in the gcDN (C). Strength product conditional on degree product (D). Edge weight conditional on degree product (E) and on strength product (F).

of its end nodes in both networks gcDN, gcMN with roughly power law dependences of the form . The scaling exponents are .

Clustering coefficient. A local property of the network structure is the tendency of the neighbors of a node i to be interconnected, which is characterized by the (topological) clustering coefficient , where is the number of triangles around the node i. (Watts and Strogatz, 1998). The average clustering coefficient as a function of degree is presented in Figure 6A for the networks gcMN and gcDN. For the gcDN we do not consider the edge directions. Both curves show power law decay of the form for the degree values up to ≈40, which cover almost all the nodes. The power law dependence of the clustering coefficient around a node on its degree is observed in many empirical networks with different scaling exponents (Newman, 2003).

In Figure 6B, we present the average weighted clustering coefficient as a function of the strength (Saramäki, 2007) for both networks gcDN, gcMN. We observe an increase of the triangles' intensity around the nodes with the increase of their strength. But, it seems that the behavior of the high-strength nodes, as seen above for the , is different from that of the others.
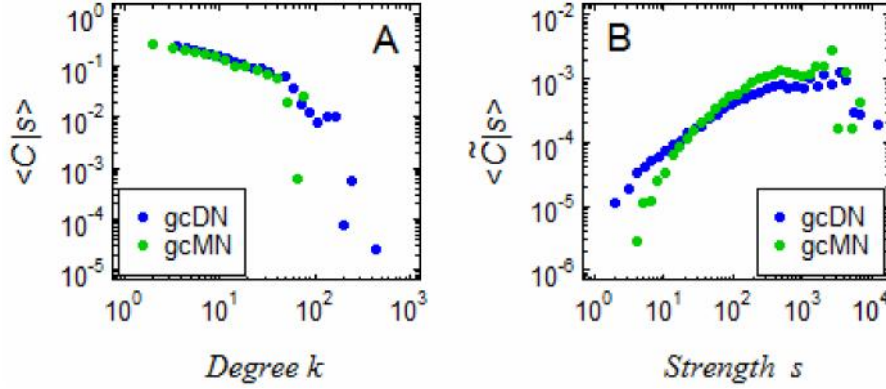
Fig.6. Topological clustering coefficient as a function of degree (A). Weighted clustering coefficient as a function of strength (B).

4. CONCLUSIONS

We have constructed two networks from mobile phone calls and SMS records of an operator in Albania. The weights of edges used to measure the strength of ties as the total number of communications during the observed time period have offered a good opportunity in exploring the mobile phone users' interactions and their relations with the topological properties of the networks.

Both networks present a large heterogeneity in the node degree and strength, and in the weight of ties which is implied by their distributions. The hubs (nodes with very high degree) in the directed network are more than those in the mutual network. This suggests that the reciprocal communication criterion for the construction of the mutual network eliminates such abnormal nodes, most likely associated with costumer service numbers or telemarketers, providing a better representation of the real social interactions.

We have found that people have the tendency to form groups with others that are similar to themselves by the number of contacts, while most of them disregard the volume of communications performed by their connections. By investigating the relationship between the communications that people perform and the number of their social contacts, we have found that people with large social circles need to perform more communications. However, the very large number of connections does not mean a proportional increase of the volume of communications. The individuals with many contacts dedicate averagely less communications to each of them than those with few connections. In addition, there exists a strong diversity in the way people schedule and manage the attention to their social relationships. The strength of a tie between two individuals does not depend on the absolute number of their social contacts, but it is positively

correlated with the number of communications they perform.

Albanian mobile phone communication networks share many common features with other similar networks studied before. Specifically, our results for the mutual network are in line with those reported by Onnela *et. al,* (2007b), who analyzed a large mutual network based on the data from an operator of an undisclosed European country with all mobile phone calls between 7 million people during a period of 18 weeks.

Li, *et. al.*(2014) studied both directed and mutual networks, similar with those we analyzed in this paper, using a dataset with detailed call records of more than nine million different mobile phone numbers from one mobile operator in Shanghai. The observations for directed networks are qualitatively similar. They compared their results of the mutual network with the aforesaid network of the European dataset and observed that both networks share many common topological features, but also several discrepancies. The mutual network of Shanghai dataset has a different broadness of degree distribution and it is not observed a dependence of the topological clustering coefficient on the node degree as observed in the European and Albanian network.

Our presented study is the first attempt in exploring and analyzing the topology and interactions' features of mobile phone communication networks with Albanian users.

REFERENCES

Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. 2004. The architecture of complex weighted networks. *PNAS USA,* **101**, (**11**): 3747-3752.

Boccaletti S, Latora V., Moreno Y, Chavez M, Hwang D-U. 2006. Complex networks: Structure and dynamics. *Physics Reports* **424**, 175-308.

Christakis N, Fowler J. 2007. The spread of obesity in a large social network over 32 years. *New England journal of medicine* **357**, 370–379.

Clauset A, Shalizi CR, Newman MEJ. 2009. Power-law distributions in empirical data. *SIAM Review*, **51**, (**4**): 661–703.

Giles J. 2012. Computational social science: making the links. *Nature* **488**, 448.

Granovetter M. 1983. The Strength of Weak Ties: A Network Theory Revisited. *Wiley, Sociological Theory*, **1**. 201-233.

Lazer D, Pentland A, Adamic L, Aral S, Barabási AL, Brewer N, Christakis NA, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alsyne M. 2009. Computational Social Science. *Science* **323**, 721–723.

Lee SH, Kim P-J, Jeong H. 2009. Statistical properties of sampled networks. arxiv:cond-mat/0505232v4.

Lewis K, Kaufman J, Gonzacles M, Wimmer A, Christakis NA. 2008. Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks* **30**, 330–342.

Li M-X, Jiang Z-Q, Xie W-X, Miccicchè S, Tumminello M, Zhou W-X, Mantegna RN. 2014. A comparative analysis of the statistical properties of large mobile phone calling networks. *Scientific Reports* **4**, 5132; DOI:10.1038/srep05132.

Mollica K, Gray B, Trevino L. 2003. Racial Homophily and Its Persistence in Newcomers' Social Networks. *Organization Science* **14**, 123–136.

Newman MEJ. 2002. Assortative mixing in networks. *Physical Review* **89**, 208701.

Newman MEJ. 2003. The Structure and Function of Complex Networks. *SIAM Review*, **45(2)**, 167–256.

Newman MEJ. 2006. Power laws, Pareto distributions and Zipf's law. arXiv:cond-mat/0412004.

Noka (Jani) E. 2014. Vertex degree distribution in mobile call networks. *Proceeding's book of SPNA 2014*, 76-81.

Onnela J, Saramäki J, Kertész J, Kaski K. 2005. Intensity and coherence of motifs in weighted complex networks. *Physical Review E* **71**, 065103.

Onnela JP, Saramäki J, Hyvöven J, Szabó G, Argollo de Menezes M, Kaski K, Barabási AL. 2007a. Structure and Tie Strengths in Mobile Communication Networks. *PNAS USA*, **104** (**18**)**:**, 7332-7336.

Onnela JP, Saramäki J, Hyvöven J, Szabó G, Argollo de Menezes M, Kaski K, Barabási AL, Kertész J. 2007b. Analysis of a large-scale weighted network of one-to-one human communication. *New Journal of Physics* **9**, 179.

Pastor-Satorras R, Vázquez A, Vespignani A. 2001. Dynamical and Correlation Properties of the Internet. *Physical Review Letters*, **87,** 258701-1.

Reed WJ, Jorgensen M. 2004. *The double Pareto-lognormal distribution - A new parametric model for size distribution. Com. Stats - Theory & Methods*, **33**, (**8**): 1733-1753.

Saramäki J, Kivelä M, Onnela J-P, Kaski K, Kertész J. 2007. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review* E **75**, 027105.

Watts DJ. Strogatz SH. 1998. Collective dynamics in 'small-world' networks. *Nature* **393**, 440–442.

Watts DJ. 2007. A twenty-first century science. *Nature* **445**, 489.

BIFURCATION ANALYSIS AS A USEFUL TOOL TO FISHERY MANAGEMENT

Valentina SHEHU
Department of Mathematics, Faculty of Natural Sciences,
University of Tirana, Albania
Alma SPAHO
Department of Applied Mathematics, Faculty of Natural Sciences,
University of Tirana, Albania

ABSTRACT

Bifurcation analysis is a key tool for the analysis of dynamic systems in general and nonlinear systems in particular. If a parameter of a dynamic system changes, qualitative behavior of the system will change. New fixed points might emerge or current fixed points might disappear, or their stability properties may change. These sorts of qualitative changes in a dynamic system are called bifurcations, and the values of the parameters at which these changes take place are called bifurcation points. Bifurcation analysis also enables us to qualitatively estimate the behavior of trajectories without utilizing the analytical solution of the underlying differential equations. Most common types of bifurcations encountered in dynamical systems are saddle-node, transcritical and pitchfork bifurcation. In the present paper some of the types of bifurcations in one-dimensional dynamical systems and application of these bifurcations in population dynamics are discussed. The fish available for human consumption comes either from the ocean or the sea. On one hand, the natural supply cannot satisfy the human needs. On the other hand, cost of fish harvesting is increasing, making aquaculture an important source for fish supply. As Albania is rich in water resources, the market demand for fish could be easily met. Aquaculture production is playing an increasing role in satisfying the demand for human consumption of fish and fishery products. Once aquaculture production is increased, the endangered species would be safe, because interventions in the rearing process to enhance production are involved. The logistic growth model is used for population growth of fish, and the two following harvesting strategies are considered: i) constant and, ii) proportional harvesting. In each process, the optimal amount of fish harvested to protect the population from extinction was estimated. The results reported that harvesting in amount or rate higher than the bifurcation point leads to the extinction of population. The results are a means to address the growth of fish population and reduce of repopulation's costs.
Keywords: fixed point, stability, saddle- node, harvesting, logistic growth model

1. INTRODUCTION

Bifurcation analysis of dynamical systems helps understand the dynamic behavior of the systems and find out possible dependence of the system's behavior upon parameter values by reporting on the occurrence and changes in stability of fixed points, and helping to model these changes and transitions from stable to unstable case or vice versa as some parameters change. In addition, bifurcation analysis enables us to qualitatively estimate the behavior of trajectories without utilizing the solution of the underlying differential or difference equations. This is achieved by numerically approximating equilibrium solutions and their stability, even for problems that do not have analytic solutions.

Bifurcation is the change in number of fixed points or periodic orbits, or in their stability properties, as a parameter varies. A bifurcation point is the value of parameter at which the change occurs. Bifurcations are classified according to how stability is lost. All types of bifurcations here reported are local bifurcations, because only the behavior of a dynamical system in the neighborhood of a single fixed point is affected.

Population dynamics is of great interest for the commercial harvesting industry and biology, ecology and economics etc. Mathematical models have been used widely to estimate the population dynamics of animals and humans. In recent years, the use of mathematical models has been extended to agriculture sector especially in cattle farming to ensure continuous and optimum supply. The logistic growth model in term of harvesting has been used to study the fishery farming (Laham *et al*., 2012). The most important for successful management of harvested populations is that harvesting strategies are sustainable, not leading to instabilities or extinctions and produces great results for the year with little variation between the years (Aanes *et al*., 2002). Therefore, it can satisfy the market demand throughout the year.

The fish available for human consumption comes either from the ocean or the sea. Unfortunately, the natural supply cannot satisfy the human needs. On the other hand, cost of fish harvesting is increasing, making aquaculture an important source for fish supply. Aquaculture production is playing an increasing role in satisfying demand for human consumption of fish and fishery products. Once aquaculture production is increased, the endangered species would be safe, as interventions in the rearing process to enhance production are involved.

As Albania is rich in water resources, the market demand for fish could be easily met. In recent years, fish farming has developed in Albania and aquaculture production constituted 20.6% of total fish caught in 2013, around 14% in 2010 and around 1% in 2001 (http://www.instat.gov.al/media/154623/tb19.xlsx).

The use of mathematical models in fishery harvesting helps the aquaculturists

to estimate the fish population for a given period. In addition, the models provide effective solutions with regard to the consumer's demand.

In the present paper, the saddle-node, transcritical and pitchfork bifurcations that occur in dynamical systems is discussed and the bifurcation analysis to harvest models is applied using the logistic growth model and the two following harvesting strategies: i) constant harvesting and, ii) proportional harvesting; to estimate the optimum quantity of fish for harvesting that can ensure the continuous supply for each strategy and, to compare the results obtained between strategies.

2. ONE DIMENSIONAL BIFURCATION

Consider a dynamical system that depends on one parameter, h:

$$\dot{x} = f(x, h) \qquad\qquad x(0) = x_0 \qquad\qquad (1)$$

When a change in a parameter results in a qualitative change in the dynamics of a nonlinear system, the system is said to have gone through a bifurcation. Bifurcations are classified by the way in which the fixed points of the function f change their number, location, form, and stability.

This paper discusses three common types of bifurcations: saddle-node, transcritical and pitchfork bifurcations (Strogatz, 1994; Lorenz, 1993; Shone, 2002; Teschl, 2004).

In a saddle-node bifurcation, as the bifurcation parameter passes through the bifurcation point, two fixed points disappear, so there are no fixed points afterwards. Before they disappear, one of the two fixed points is stable and the other is unstable.

If the differential equation (1) satisfies the equations:

$$f(x_0, h_0) = 0$$

$$\frac{\partial f}{\partial x}(x_0, h_0) = 0$$

$$\frac{\partial^2 f}{\partial x^2}(x_0, h_0) \neq 0$$

and
$$\frac{\partial f}{\partial h}(x_0, h_0) \neq 0$$

then this equation has a saddle node bifurcation at $h = h_0$.

The key example of a saddle node bifurcation is provided by the differential equation:

$$\dot{x} = h - x^2 \qquad\qquad (2)$$

Clearly, we have $f(0,0) = 0$ and $\dfrac{\partial f}{\partial x}(0,0) = 0$. Fixed points of this equation are given by $h = x^2$, corresponding to a parabola in the xh plane. For $h < 0$, the equation has no fixed points and $\dot{x}$ is always negative. For $h > 0$, there exist two fixed points, one stable and one unstable.

In dynamical systems, a bifurcation diagram shows the possible long-term values (fixed points or periodic orbits) of a system as a function of a bifurcation parameter in the system. The stable fixed points are represented by a green line and unstable ones by a red line. The bifurcation diagram for the saddle-node bifurcation indicates again that the two fixed points collide and annihilate at $h = 0$ and there are no fixed point for $h < 0$ (Figure 1).

Fig.1: Bifurcation diagram for the saddle-node bifurcation.
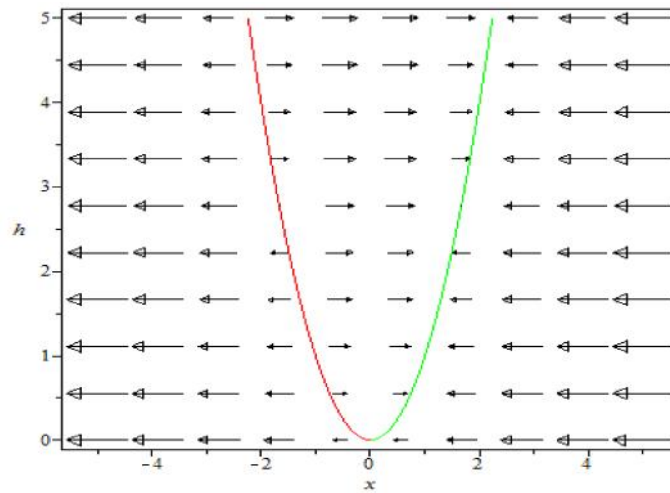
In a transcritical bifurcation, there are two fixed points, one stable and one unstable. When the bifurcation point is passed, there is an exchange of stability; the unstable fixed point becomes stable and the stable one becomes unstable.

If the differential equation (1) satisfies the equations:

$$f(x_0, h_0) = 0$$

$$\frac{\partial f}{\partial x}(x_0, h_0) = 0$$

$$\frac{\partial f}{\partial h}(x_0, h_0) = 0$$

$$\frac{\partial^2 f}{\partial x^2}(x_0, h_0) \neq 0,$$

and $\dfrac{\partial^2 f}{\partial x \partial h}(x_0, h_0) \neq 0$,

then this equation has a transcritical bifurcation at $h = h_0$.

The simplest example of a transcritical bifurcation is provided by the following differential equation:

$$\dot{x} = hx - x^2 \qquad\qquad (3)$$

Clearly, we have $f(0,0) = 0$ and $\dfrac{\partial f}{\partial x}(0,0) = 0$. The fixed points of this equation are $x = 0$ and $x = h$, corresponding to two lines in the xh plane. For $h < 0$, the equation has two fixed points, $x = 0$ is stable and $x = h$ is unstable. The two fixed points collapse at $h = 0$ and for $h > 0$, $x = 0$ is unstable and $x = h$ is stable. An exchange of stability has occurred at $h = 0$.

The bifurcation diagram for the transcritical bifurcation is depicted in figure 2.
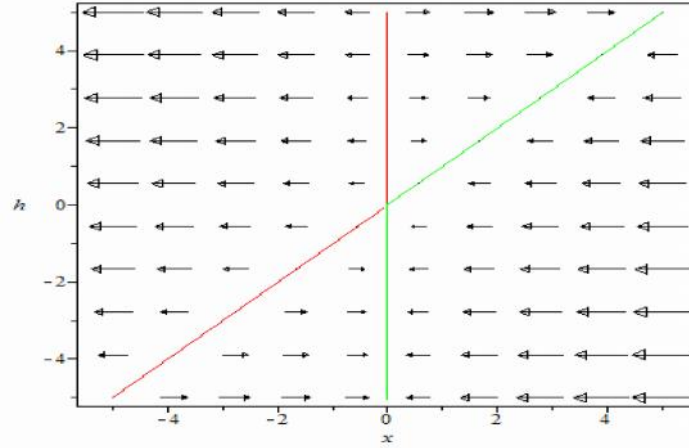

Fig. 2: Bifurcation diagram for the transcritical bifurcation.

In a pitchfork bifurcation, there are two stable fixed points separated by an unstable one. Once bifurcation point is passed, only one stable fixed point remains. This type of bifurcation is called supercritical pitchfork bifurcation.

There is also a subcritical pitchfork bifurcation, in which there are two unstable fixed points separated by a stable one, until the bifurcation point is passed. Then there is only one unstable fixed point.

If the differential equation (1) satisfies the conditions:

$$f(x_0, h_0) = 0$$

$$\frac{\partial f}{\partial x}(x_0, h_0) = 0$$

$$\frac{\partial f}{\partial h}(x_0, h_0) = 0$$

$$\frac{\partial^2 f}{\partial x^2}(x_0, h_0) \neq 0$$

$$\frac{\partial^2 f}{\partial x \partial h}(x_0, h_0) \neq 0$$

and $\dfrac{\partial^3 f}{\partial x^3}(x_0, h_0) \neq 0$,

then this equation has a pitchfork bifurcation at $h = h_0$.

As an example of a pitchfork bifurcation consider the following differential equation:

$$\dot{x} = hx - x^3 \qquad\qquad (4)$$

Clearly, we have $f(0,0) = 0$ and $\dfrac{\partial f}{\partial x}(0,0) = 0$. The fixed points of this equation are given by $x = 0$ and $h = x^2$, corresponding to a line and a parabola in the xh plane. For negative values of h, the equation has one stable fixed point at $x = 0$. For $h > 0$, $x = 0$ is a unstable fixed point and the two new fixed points have been created at $h = 0$, one in each branch of the parabola $h = x^2$.

The bifurcation diagram for the supercritic pitchfork bifurcation is depicted in figure 3.
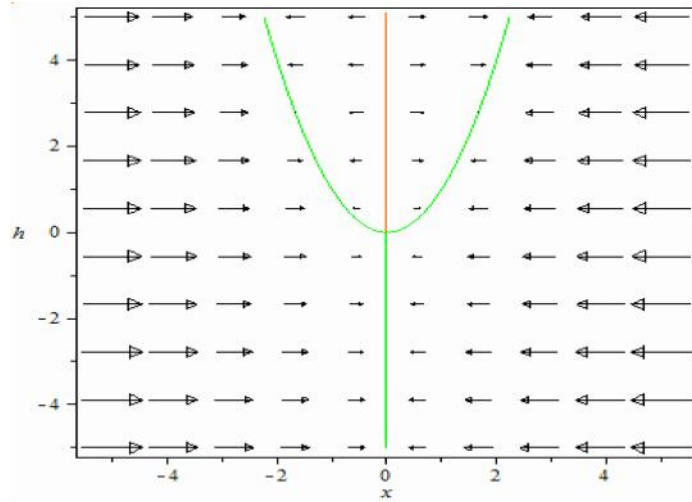
Fig. 3: Bifurcation diagram for the supercritic pitchfork bifurcation.

Next section discusses the application of these types of bifurcation to fishery management.

3. APPLICATION OF BIFURCATIONS: FISHERY MATHEMATICAL MODELS

To illustrate the bifurcation analysis in a real world situation we are going to analyze some simple fishery management models. Consider the logistic growth equation to model a fish population in the absence of fishing:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{M}\right) \tag{5}$$

where P is the size of the population, r is the growth rate due to reproduction and M is the carrying capacity of the environment. Now considering modeling the population and harvesting some of the population using some common harvesting strategies: constant harvesting and proportional harvesting. Constant harvesting is where a fixed numbers of fish were removed each year, while periodic harvesting is usually thought of a sequence of periodic closure and openings of different fishing grounds (Idels and Wang, 2008; Laham *et al.*, 2012). In proportional harvesting, the quantity harvested is proportional to the population. Harvesting has been considered a factor of stabilization, destabilization, and improvement of mean population levels, induced fluctuations, and control of non-native predators (Michel, 2007).

Constant harvesting

One of the simplest methods is to establish a set limit for harvesting. We assume that the dynamics of the population satisfies the logistic growth model (5) and that a constant harvesting, h, is added for removing a constant number of the fish over a given time interval.

The mathematical model becomes:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{M}\right) - h \qquad \qquad P(0) = p_0 \qquad \qquad (6)$$

The solution of this ordinary differential equation obtained using MAPLE is difficult to analyse, so we turn to a geometric analysis of the model.

The fixed points, P^*, are simply the solution of the equation $rP^*(1 - P^*/M) = h$.

The model (6) has two roots: $P_{1,2}^* = \frac{1}{2}\left(M \pm \sqrt{M^2 - \frac{4hM}{r}}\right)$. The number of

fixed points depends on the sign of the term M^2 - 4hM/r. There are: two fixed points if M^2 - 4hM/r > 0 or equivalently $0 < h < rM/4$; one fixed point if M^2 - 4hM/r = 0 or equivalently $h = rM/4$; and no fixed point if M^2 - 4hM/r < 0 or equivalently $h > rM/4$.

To show the above results, conditioned by the unavailability of any real and reliable data, consider a lake or pond with a rare fish and suppose that the rate of population growth is $r = 0.7$, carrying capacity is M = 1000. The fixed points for model (5), where $h = 0$, are $P^* = 0$ and $P^* = M = 1000$.

The bifurcation diagram for the constant harvesting logistic model is depicted in figure 4.
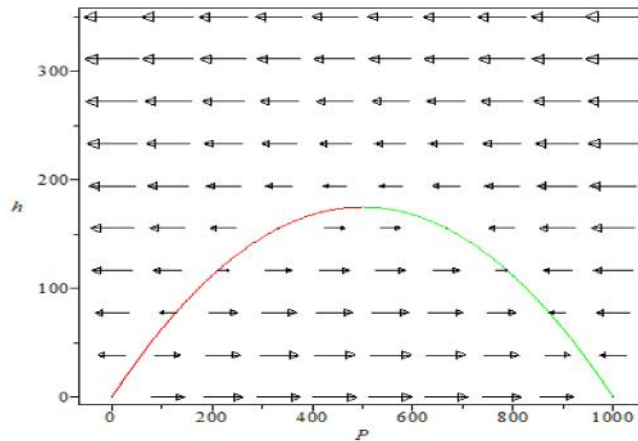

Fig. 4: Bifurcation diagram for the constant harvesting.

As the harvesting increases, the two fixed point move closer to each other with the lower fixed point remaining unstable and the upper fixed point remaining stable. As *h* moves toward the maximum growth rate of the logistic growth equation, the two fixed points coalesce at $P^* = M/2 = 500$. When $h > rM/4 = 175$, there are no fixed point, and the model shows that the population always goes extinct. This model shows a classic example of a *saddle node bifurcation*. The bifurcation values is at $h = rM/4 = 175$.

The purpose of simple mathematical models applied to complex problems is to offer some insight. Here, the results reported that overfishing (in the model $h > rM/4$) during one year can potentially result in a sudden collapse of the fish catch in subsequent years, so that governments need to be particularly cautious when contemplating increases in fishing quotas.

Proportional harvesting

Another common form of harvesting is when one puts in a constant effort to harvest. In this case, the quantity harvested is proportional to the population. Thus, the mathematical model is written:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{M}\right) - hP \qquad\qquad P(0) = p_0 \qquad (7)$$

where again *r* is the growth rate, M is the carrying capacity with no harvesting and now *h* is the proportional rate of harvesting.

Algebraic solution is complex and harder to interpret, thus we again turn to the geometric analysis of the model. The fixed points of (7) are the solution of the equation: $rP^*\left(1 - \frac{P^*}{M}\right) = hP^*$, that is, and $P^* = \frac{M(r-h)}{r}$. The extinction fixed point, $P^* = 0$, is unstable for values of $h < r$. As *h* increases, the larger equilibrium (carrying capacity, M) shrink, but it remains stable for $h < r$. For $h = 0$, $P^* = M = 1000$.

The bifurcation diagram for proportional harvesting logistic model is depicted in figure 5.
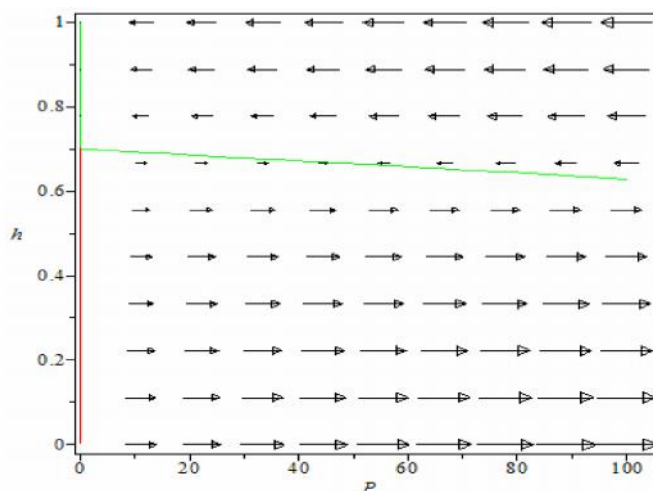
Fig. 5: Bifurcation diagram for proportional harvesting.

As the harvesting increases, the nontrivial fixed point move closer to extinction fixed point. As h moves toward the growth rate (0.7), the nonzero fixed point fades to zero, which implies there is extinction. This is comprehensible due to the harvesting rate approaching the growth rate. When $h > r$ (= 0.7), the rate of harvesting exceeds the reproduction rate and extinction necessarily follows. This model shows a classic example of a *transcritical* bifurcation. The bifurcation point is $h = r$ (= 0.7).

The results reported that overfishing ($h > r$) in one year can potentially extinct the fish in the lake or pond. Consequently, the bodies involved in the area need to be particularly cautious when contemplating fishing quotas.

4. CONCLUSIONS

Differential equations containing parameters with approximately known values are commonly involved in practical applications. Consequently, investigating the solutions' behavior and their dependence on the parameters is very important, as a slight variation in a parameter might have significant impact on the solution.

The three common types of bifurcations here reported are: i) saddle-node bifurcation, ii) supercritical pitchfork bifurcation and, iii) transcritical bifurcation. In each case, the bifurcation value is $h = 0$. These types of bifurcations also occur in higher-dimensional dynamical systems.

Sustainable harvesting strategies and successful management of harvested populations help the fishery industry to grow both qualitatively and quantitatively. However, some little variation in fish population between the years could still

be reported. Regarding the constant harvesting strategy, the fish population does not have enough time to recover if the constant harvesting is greater than the bifurcation value $h = rM/4$. Regarding the proportional harvesting strategy, the fish population will extinct if the proportional rate of harvesting is greater than the growth rate of the population, $h > r$.

The results are a means to address the growth of fish population and reduce of fish repopulation's costs. In addition, the annual overfishing has unavoidably impacts fishing in the subsequent years. Here, an appropriate national policy on fishing quotas would be crucial.

Appropriate fishery harvesting strategies would satisfy the market demand. Commercializing the aquaculture could be a good alternative.

However, some limitations could be here encountered: i) theoretical results, ii) unreal parameters of logistic model for fish harvesting is reported.

The application of bifurcation analysis to fish harvesting in real data is a means to address further researches in the area. In addition, other practical applications of bifurcations such as one-dimensional or higher dimensional dynamical systems must further.

REFERENCES

Aanes S, Engen S, Saethe B-E, Willerbrand T, Marcstram V. 2002. Sustainable harvesting strategies of willow ptarmigan in a fluctuating environment. *Ecological Applications* **12**: 281-290.

Idels LV, Wang M. 2008. Harvesting fisheries management strategies with modified effort function. *International Journal Modelling, Identification and Control* **3**: 83-87.

INSTAT (database): http://www.instat.gov.al/media/154623/tb19.xlsx

Laham MF, Krishnarajah IS, Shariff JM. 2012. Fish harvesting management strategies using logistic growth model. *Sains Malaysiana* **41(2)**: 171–177.

Lorenz H-W. 1993. Nonlinear dynamical economics and chaotic motion. Berlin: Springer-Verlag. 80-87.

Michel IDSC. 2007. Harvesting induced fluctuations: insights from a threshold management policy. *Mathematical Biosciences* **205**: 77-82.

Shone R. 2002. Economic dynamics phase diagrams and their economic application. Cambridge University Press 287-293.

Strogatz SH.1994. Nonlinear dynamics and chaos with applications to physics, biology, chemistry and engineering. Cambridge MA: Perseus 44-60.

Teschl G. 2004. Ordinary Differential Equations and Dynamical Systems 198-200.

ESTIMATION OF THE ION AND AEROSOL CONCENTRATIONS AND THEIR RELATIONSHIP WITH METEOROLOGICAL PARAMETERS IN A SEASHORE SITE

Floran VILA
Department of Physics, Faculty of Natural Sciences, University of Tirana, Albania
Florian MANDIJA
Department of Physics, Faculty of Natural Sciences, University "Luigj Gurakuqi", Shkrodra, Albania
Fatos YLLI
Institute of Applied Nuclear Physics, University of Tirana, Albania
Dhurata PREMTI
Industrial Chemistry Department, Faculty of Natural Sciences, University of Tirana, Albania

ABSTRACT

Cluster ions and aerosols are atmospheric particles, which participate and have a great impact on several global processes. On other hand, both these particles interact with each other via recombination and attachment processes. Monitoring the variation of their concentrations gives valuable information about their interactions and the impact on atmospheric processes that these particles take part. The present paper provides information about the variation of particle number concentrations of cluster ions and aerosol particles of sub-micrometric and super-micrometric size. Investigation was carried out from 2009-2011 at a site located in the Adriatic seashore and the results reported high values of cluster ions and aerosols. The source is anthropogenic activities, sea salt and other particles of long-range transport.
Keywords, Ion concentrations, aerosol concentrations, seashore site, anthropogenic activities

1. INTRODUCTION

The major part of the lower troposphere consists of atmospheric particles such as cluster ions and aerosol. The formation and growth of atmospheric aerosols depend on several steps, namely nucleation, initial steps of growth

and subsequent – mainly condensational – growth. The main source is anthropogenic activity, but they are also formed naturally in seas, deserts, and forests. Recent data report that production of new aerosol particles by nucleation and subsequent growth by condensation of organic vapours emitted by trees is a frequent phenomenon that occurs in most atmospheric environments (Kulmala *et al.*, 2004).

Cluster ions differ from aerosol particles primarily by their smaller size or greater electrical mobility (Horrak, 2001). Other differences between these two groups relate to their electric charge, spatial distribution, vertical profiles, size distribution functions, etc. Atmospheric ions, according to their electrical mobilities may be divided into three main categories (table 1) (Horrak *et al.*, 2003).

Table 1. Categories of atmospheric ions

	Diameter Range nm	Mobility Range, $cm^2 V^{-1} s^{-1}$
Small ions	0.36 - 1.6	1.3 – 0.5
Intermediate ions	1.6 - 7.4	0.5 - 0.034
Large ions	7.4 - 80	0.034 - 0.00041

Aerosol particles, according to their size, could be divided into the following groups: i) ultrafine (3 – 90 nm), ii) fine (90 – 1000 nm) and, iii) coarse (> 1000 nm). Ultrafine mode may be divided also into two other modes; nucleation and Atiken (on threshold 20 nm). Unlike atmospheric ions, aerosol particles can be neutral or single/multi charged (Mandija, 2011-a). Table 1 reports that two large categories of atmospheric ions (intermediate and large ions) are charged ultrafine aerosol particles. Thus, the traditional border between cluster ions and charged aerosol particles is 1.6 nm (Tammet, 1995).

Both atmospheric ions and aerosols play important roles in atmospheric processes (Flagan, 1998). Air ions make the air fresh, helping persons who suffer from asthma or other respiratory problems. Meanwhile atmospheric aerosols, "swallowing" air ions, play an inverse role on environmental concern (Rose *et al.*, 2013).

Concentrations of atmospheric ions and aerosols are governed by balance equation. This equation in its simplest form is presented by eq.1 (Salm and Tamm, 2011).

$$\frac{dn}{dt} = q - \Gamma n_- n_+ - S_+ n_+ N_+ - S_- n_- N_- - S_0 (n_- + n_+) \tag{1}$$

where: $n_{\pm}$ is the concentrations of ions of different polarities

$N_{\pm,0}$ is the concentrations of charged and neutral aerosols

г and ѕ are their respective recombination and attachment coefficients.

Larger ions can swallow more cluster ions, but because of their minor concentrations, there are small aerosols that influence more of the ion concentrations.

The present investigation was carried out at Velipoja beach which is a tourist location on the Adriatic seashore. Air studies are of great importance, as they help evaluate air quality (Mandija *et al.*, 2010; Mandija, 2011-b).

2. MATERIAL AND METHODS

Measurement of atmospheric ion and aerosol concentrations was carried out at Velipoja, a tourist location on the Adriatic seashore in Albania. Figure 1 depicts site location 200 m from the seashore, in the middle of a residential area.

Investigation was carried out in spring, summer and autumn (April-October). In summer, anthropogenic activity intensifies along with the population, as several tourist-commercial activities take place. In spring and autumn, anthropogenic activities are very rare.

Measurements were carried from 2009 to 2011. As only measurement results regarding fair-weather conditions were analyzed, the impact of precipitations and strong winds on atmospheric particle concentrations were excluded.

Fig.1 Monitoring site, Velipoja Beach, Albania (from Google Maps).

The Air Ion Counters IC-1000 and Andes ITC-201A were used to measure the concentrations of atmospheric cluster ions (d < 1.6 nm). Operation principle of this instrument is the appliance of electric field between three parallel conductive plates. The middle plate collects air ions which enter into the intern area of the AIC. Atmospheric aerosol concentration was measured using a Handheld Particle Counter, model Handilaz mini 301. This instrument measures aerosol concentrations divided into three channels; in the size range over 0.3

µm. Thresholds between channels varied from 0.5 and 5 µm. Aerosols on the first channel belong to fine mode, whilst those on the third channel belong to coarse group. Aerosols on second channel belonged both modes, but more to the fine mode (because their size distribution). Operation principle of this instrument is laser light scattering.

3. RESULTS

3.1 Ion concentrations

Table 2 reports the average values of meteorological parameters during the monitoring process and temperature, relative humidity, atmospheric pressure and wind speed had normal values throughout the monitoring process. Usually, in summer temperature is higher and wind speed has lower values

Table 2. Average values of meteorological parameters

	Summer	Spring and Autumn
Temperature (^{o}C)	28.6	19.6
Relative humidity (%)	62.3	57.3
Atmospheric pressure (hPa)	1205.1	1027.3
Wind speed (km/h)	1.2	3.7

Table 3 reports the average values and the standard deviations of ion concentrations of both polarities.

Table 3. Seasonal-averaged ion concentrations

	Summer	Spring and Autumn
Positive (cm^{-3})	1752.0±843.0	1145.8±119.6
Negative (cm^{-3})	1550.0±665.1	487.5±100.3
Total (cm^{-3})	3302.0±1482.8	1633.3±182.6
Unipolarity factor	1.1±0.2	2.5±0.7

The average values of the concentrations of both positive and negative ions in summer were higher (positive 1.5 and negative 3.1 times) than in spring and autumn. Since the anthropogenic activities, like traffic, heating-cooking, etc. are extremely higher in summer due to the tourism air ion concentrations are higher (Yu, 2002). In addition to the anthropogenic activity, vegetation in the surrounding forest has a great impact (Jayaratne *et al.*, 2011). This fact implies the hypothesis that the principal sources of air ions in Velipoja beach are anthropogenic activities and vegetation. Moreover, the dry summer season also enhances radon gas emanation from the ground surface, increasing its concentration in the regions near the ground (Hosoda *et al.*, 2007). Both these

factors have a renewed presence in this region during summer seasons.

Table 3 and 4 report that the meteorological parameters had a great impact on the ion concentrations. As the temperature is higher in summer, ion concentrations are higher. So, we can suppose that temperature influences the increase of ion concentration (Luts, 1998) indirectly due to the augmentation of radon exhalation (Porstendorfer, 1994; Hopke and Ramamurthi, 1998; Katoka *et al*., 1998). Regarding the wind, its speed affects directly the ion concentration via dispersion. These hypotheses are confirmed by the correlation coefficients between ion concentrations, temperature and wind speed. Table 4 presents the Pearson correlation coefficients among the ion concentrations and the principal meteorological parameters.

Table 4. Correlation coefficients among aerosol concentrations and meteorological parameters

Meteorological parameters	Ion concentration
Temperature	0.26 (0.04)
Relative humidity	-0.11 (0.38)
Atmospheric pressure	-0.13 (0.37)
Wind speed	-0.41 (0.01)

Significant levels of the correlation coefficients are given by the values in brackets. The correlation coefficients among total ion concentrations and tempreature are statistically significant ($p < 0.05$). Statistically significant correlation coefficients were found also in the case of ion concentrations and wind speed values. Correlation coefficients between ion concentrations and relative humidity and atmospheric pressure were even lower, respectively -0.11 and -0.13. Their significance level is higher than 0.05, showing that their correlations are not statistically significant. These values indicate that these meteorological parameters play a minor role on ion concentrations. The aforementioned data are based on the statistical interpretation of the measurement data.

3.2. Aerosol concentrations

The table 5 reports the average values and their standard deviations of the aerosol concentrations of the three modes. Major emissions occur during the summer season. Aerosol emissions belong primarily to fine mode. The average concentrations of fine aerosols during summer season were more than 1.5 times higher than during spring and autumn seasons.

Summer marks intensive combustion processes, construction activities, biomass burning and traffic (Jayaratne *et al*., 2010). These processes are

predominant aerosol sources in summer; construction activities and biomass burning occur also during other seasons. The presence of coarse aerosols is related to local wind re-suspensions processes.

Table 5. Seasonal-averaged aerosol concentrations

	Summer	**Spring and Autumn**
Fine (10^6 m^{-3})	5.2±0.9	5.0±0.7
Mixed (10^6 m^{-3})	1.1±0.6	0.9±0.5
Coarse (10^4 m^{-3})	10.4±8.6	1.9±1.3

To estimate the effect of meteorological parameters on aerosol concentrations, their Pearson correlation coefficients were determined (table 6).

Table 6. Correlation coefficients among aerosol concentrations and meteorological parameters

	Fine	**Mixed**	**Coarse**
Temperature	0.06 (0.63)	0.14 (0.27)	0.47 (0.01)
Relative humidity	0.80 (0.01)	0.73 (0.01)	0.22 (0.08)
Atmospheric pressure	0.20 (0.16)	0.52 (0.01)	0.01 (0.99)
Wind speed	0.16 (0.22)	0.28 (0.03)	–0.18 (0.17)

Relative humidity has a greater impact on aerosol concentrations, especially those of the fine and intermediate modes than the other meteorological parameters. The instrument here used measures also liquid aerosols which concentrations increase under the high relative humidity conditions. On the other hand, under high temperature conditions earth surface undergoes on low soil moisture values. Under these conditions, re-suspension processes are favored, and aerosol concentrations of coarse mode are enhanced. Statistically, no significant correlation was obtained in the case of the wind speed and coarse aerosol concentrations. This was due to the limited number of available measurement data.

3.3. Binom ion-aerosol

Here, the correlation between ion and aerosol concentrations is investigated. It is a well known the fact that fine aerosols (because of their major number) influence more on the reduction of the concentrations of cluster ions, expressed by eq.1 (Dhanorkar and Kamra, 1992; Reischl *et al*., 1996; Makela *et al*., 2001).

To estimate this contribution in terms of correlation coefficients, should be determined these coefficients (Hoppel and Frick, 1986). The correlation coefficient between the concentrations of cluster ions and fine aerosol particles

is -0.21 (0.15)—statistically insignificant. A correlation of -0.31 (0.03) was found in the case of the mixed aerosols. Coarse aerosols were weakly correlated with cluster ions with correlation coefficient 0.23 (0.12). Thus, the presence of fine and intermediate aerosols at the site affects the populations of the cluster ions, reducing their concentrations.

Table 7 reports correlation coefficients among ion concentrations, aerosol concentrations (mixed mode) and meteorological parameters.

Table 7. Correlation coefficients among ion concentrations and principal influencing factors

Mixed aerosols	Temperature	Wind speed
-0.31 (0.03)	0.26 (0.04)	-0.41 (0.01)

The data in Table 7 reports that wind activity, aerosol concentrations and the temperature are of great impact for the ion concentration. The concentration of the aerosols decreases ion concentrations due to the attachment processes between them. Wind activity decreases the ion concentration, due to the dispersion process. Meanwhile, temperature acts indirectly, increasing the radon concentration and consequently the concentration of cluster ions. Anyway, for an accurate determination of the influence of meteorological parameters on the ion concentrations, the mutual relationship between ion and aerosol concentrations as well as the meteorological parameters should be taken into account. In addition, a larger set of data will enable to obtain more significantly correlation coefficients.

Figure 2 depicts ion concentrations and influencing factors in different seasons. The correlation coefficients of these two sets of data (summer and other seasons) are very high, 0.97. This is another important result, which shows that these factors undergo almost in the same direction regardless of the season.

All units of the quantities in figure 2 are expressed based on the tables 2, 3 and 5.
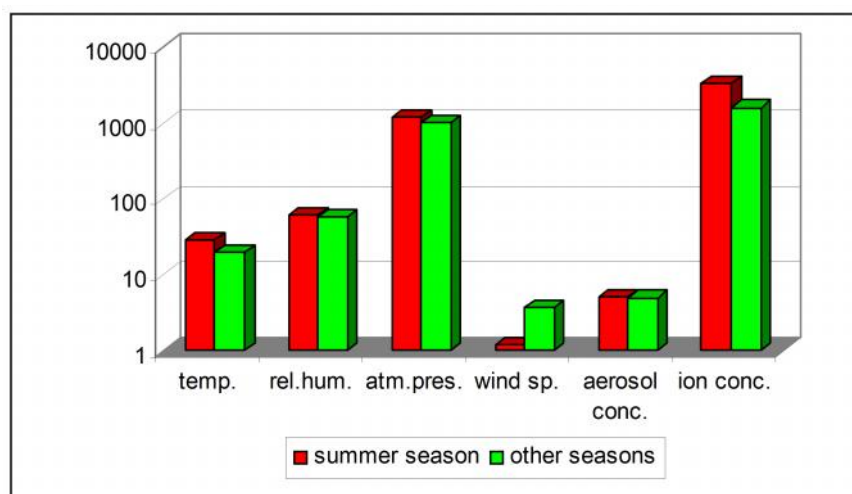
Fig. 2 Ion concentrations and influencing factors in different seasons.

4. CONCLUSIONS

The data of ion and aerosol concentrations at the Velipoja site are here reported. The average values of these concentrations indicate generally fresh air, (1000-5000 cm^{-3}). The source is anthropogenic activity (traffic, combustions and vegetation). The source for fine aerosols is also anthropogenic activities (traffic, combustion and construction). Re-suspension processes mostly affect the coarse mode.

In addition to anthropogenic activity, the meteorological parameters have a great impact on these concentrations. Temperature has a greater impact on ion concentration. Relative humidity has a great impact on the concentrations of fine and mixed aerosols. Wind activity has a great impact on the concentrations of the coarse aerosols and the air ions. Mixed aerosols reduce the concentrations of cluster ions.

Thus, principal sources and eliminators of cluster ions on a seashore site are determined, within a wide range of period, based on the comparison of measurement data during summer, spring and autumn.

REFERENCES

Dhanorkar S, Kamra AK. 1992. Relation between electrical conductivity and small ions in the presence of intermediate and large ions in the lower atmosphere, *Journal of Geophysical Research*, **97**, 20345–20360.

Flagan RC. 1998, History of electrical aerosol measurements, *Aerosol Science and Technology*, 28:4, 301-380.

Hopke PK, Ramamurthi M. 1988, Production of ultrafine particles by radon radiolysis, *Journal of Aerosol Science*, **19**, 1323–1325.

Hoppel WA, Frick GM. 1986. Ion-aerosol attachment coefficients and the steady-state charge distribution on aerosols in a bipolar ion environment, *Aerosol Science and Technology*, 5, 1–21.

Horrak U, 2001, Air ion mobility spectrum at a rural area, *Dissertationes Geophysicales Universitatis Tartuensis*, **15**, 157 pp., Tartu Univ. Press, Tartu, Estonia.

Horrak U, Salm J, Tammet H. 2003. Diurnal variation in the concentration of air ions of different mobility classes in a rural area, *Journal of Geophysical Research*, Vol. 108, No. 4653, 7-1 – 7-11.

Hosoda M, Shimo M, Sugino M, Furukawa M and Fukushi M. 2007. Effect of Soil Moisture Content on Radon and Thoron Exhalation. *Journal of Nuclear Science and Technology*, 44:4, 664-672.

Jayaratne ER, Ling X, Morawska L. 2010. Ions in motor vehicle exhaust and their dispersion near busy roads, *Atmospheric Environment*, **44(30)**; 3644 – 3650.

Jayaratne ER, Ling X, Morawska L. 2011. Role of Vegetation in Enhancing Radon Concentration and Ion Production in the Atmosphere. *Environmental Science & Technology*, 45(15), 6350-6355.

Kataoka T, 1998, Diurnal variation in radon concentration and mixing layer depths, *Boundary-Layer Meteorology* , 89, 225–250.

Kulmala M, Vehkamaki H, Petaja T, Dal Maso M, Lauri A, Kerminen VM, Birmili W, McMurry PH. 2004. Formation and growth rates of ultrafine atmospheric particles: A review of observations, *Journal of Aerosol Science*, 35, 143–176.

Luts A. 1998. Temperature variation of the evolution of positive small air ions at constant relative humidity, *Journal of Atmospheric and Solar-Terrestrial Physics*, **60**, 1739–1750.

Makelä JM, Salm J, Smirnov VV, Koponen I, Paatero J, Pronin AA. 2001. Electrical charging state of fine and ultrafine particles in boreal forest air, *Journal of Aerosol Science*, **32**, S149–S150.

Mandija F, 2011-a, A method to determine the recombination and attachment coefficients of atmospheric ions and ultrafine particles in different charge levels, *RSC Advances*, 1, 142 –146.

Mandija F, Ahmentaga Sh, Vila F, 2010, Atmospheric electricity monitoring on the Adriatic seashore, *AIP Conference Proceedings*, **1203**, 547-552.

Mandija F, 2011-b, Concentrations of atmospheric ions and aerosol particles near Adriatic Sea, *Environmental Research Journal,* **4, (5-6);** 591 – 608.

Porstendorfer J. 1994. Properties and behavior of radon and thoron and their decay products in the air, *Journal of Aerosol Science*, **25**; 219 – 263.

Reischl GP, Makela JM, Karch R, Necid J. 1996. Bipolar charging of ultrafine particles in the size range below 10 nm, *Journal of Aerosol Science*, **27**; 931–949.

Rose C, Boulon J, Hervo M, Holmgren H, Asmi E, Ramonet M, Laj P, Sellegri K. 2013. Long-term observations of positive cluster ion concentration, sources and sinks at the high altitude site of the Puy de Dom. *Atmospheric Chemistry and Physics*, **13**, 14927–14975.

Salm J, Tamm E. 2011. Dependence of the Ion-Aerosol Equivalent Attachment Coefficient on the Ratio of Polar Conductivities in a Steady State. *Aerosol and Air Quality Research*, **11**; 211–217,

Tammet H. 1995. Size and mobility of nanometer particles, clusters and ions, *Journal of Aerosol Science*, **26**, 459-475.

Yu F. 2002. Chemiion evolution in motor vehicle exhaust: further evidence of its role in nanoparticle formation, *Geophysical Research Letters*, **29, (15);** 1717.

USING DATA MINING FOR EMPLOYMENT PREDICTION OF GRADUATES IN ALBANIA

Alba ÇOMO and Xheni MELO
Department of Informatics, Faculty of Natural Science, University of
Tirana, Albania
Jona MULLIRI
Department of Mathematics and Informatics, Faculty of Economics and
Agribusiness, Agricultural University of Tirana, Albania

ABSTRACT

A survey was carried out to investigate the employment opportunities of Albanian newly graduated students and the data are here reported. The data mining classification techniques such as NaiveBayesUpdatable, Logitboost and Adaboost1 algorithms were in the present investigation applied. The data collected were investigated based on: i) prediction accuracy, ii) model building time and, iii) error rate. In the end, the most appropriate algorithm was determined.
Keywords: Data mining, Classification Model, Graduates Employment, Prediction

1. INTRODUCTION

Elevated unemployment imposes significant costs on the individual and society. The present paper investigates the employment rate of Albanian newly graduated students. In Albania, unemployment rate is high and very concerning. Given the social-economic situation, such information would be relevant for the institutions involved in the area.

The quality of education system is very important for the employment of graduated students. Çela and colleagues (2015) reported that number of students that fail to graduate is high. Consequently, finding a job is hard. Given the situation, categorizing the students' profile would be very important. As this is a new and unexplored investigation area in Albania, difficulties in data collection even from a non-computerized system were encountered.

The present investigation aims at predicting the success of the graduated students via data mining techniques. In the end, the data mining technique was

determined based on the accuracy.

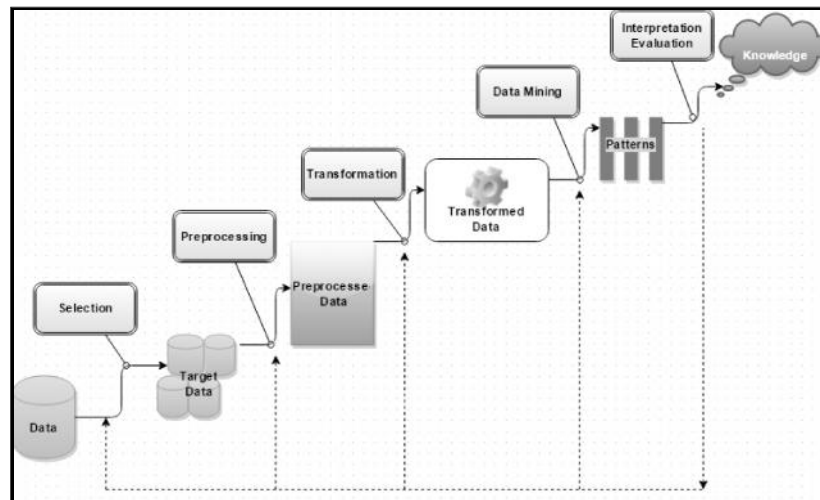Figure 1 depicts the process here carried out.

Fig. 1: Knowledge discovery process

2. MATERIALS AND METHODS

In the present investigation, the efficiency of three data mining techniques such as neutral networks, the decision trees and the Naive Bayes were compared. The data were collected from a survey sent via various social networks to a group of graduated students whose demographic data were collected along with their education and employment experience.

2.1 Literature review and definitions

Data Mining is the process that uses several tools for data analysis to find models and connection between the data, in order to make the prediction (Witten and Frank, 2005).

Regarding the academic process, data mining could be used to: i) estimate students' performance, ii) make recommendations for them, iii) estimate the know-how process, iii) prepare them to adopt and use different techniques in their job based on job market tendencies and, iv) foster intellectual pursuits in different areas. Consequently, the job market wouldn't be challenging (Romeo and Ventura, 2007; Al-Radaideh and Nagi, 2012).

The future of a country depends on its youth potential which source is the education system (Muntean *et al*., 2010).

Data mining is used to estimate the employment opportunities, unemployment

rate, job market requirements etc. In addition, it could be used by different companies when hiring. The cause-consequence relationship is a means to address education system towards job market tendencies and development practices towards multiple investments (Muntean, *et al*. 2010).

As classification models are constructed on the basis of historical data, data mining might lead to racial, social, religious discrimination etc. Consequently, accuracy would be relevant.

2.2 The study process

Zhang and colleagues (2010) said that evaluating the results of students into the job market involves certain criteria. In the present paper, data on education level have been investigated.

The present investigation provides some information about the employment opportunities of the newly graduated students in Albania. Consequently, different methods were applied and in the end the most appropriate one was determined based on prediction accuracy.

2.3. Data collection

The data were collected from a survey sent via various social networks to a group of graduated students aged between 18 and 27 in Tirana and Korça. Their demographic data were collected along with their education level, working experience, age, gender, residence job selection criteria etc. This age group was chosen as they are expected either to attend university or foster intellectual pursuit. The survey aimed to give accurate predictions on the employment possibilities of the Albanian newly graduated students. The unemployed graduated students were asked on the reasons of unemployment. The survey was sent online via social networks such as Facebook, Google+, LinkedIn, etc. Once the incomplete data were removed, the total number of the newly graduated students involved properly in the survey was 180.

The model of success of graduated students was created. Here, success of the model was evaluated based on their employment success (based on the result variable–employed– yes/no). The prediction model consists of 10 variables and the Employed class. The chosen variables are in the Table 1 reported.

Table 1 The Chosen Variables

Variables	Description	Values
Gender	Student Gender	{M,F}
Age	Student Age	18-27
Residence	Student Residence	{ A---Tirana, B---Korca}
Completed Studies	Type of studies completed or that are attending	{A –PhD, B- Bachelor, C- Master, D – High School, E- Primary School, F- Attending Bachelor/Master}
Country of Studies	Country where you have studied	{A--Albania, B--Abroad}
Type of School	The type of school attended	{A---Public, B---Private}
Program	Study Program	
Average	The last study program Average (or the average of Higher Education, if is still attending)	
Foreign Language	The number of the foreign languages you can speak	{A---1, B---2, C ---3, D--- >3}
Training	Have you attended training or supplementary courses	{A---Yes, B---No}
Employment	Your employment status	{A---Employed, B---Unemployed}

2.4. The Data mining methods

In the present investigation The *Naïve Bayes Updateable* (NBU), The *LogitBoost* (LB) and, The *AdaBoost1* (AB1) classifier were involved.

The NBU classifier is part of Naïve Bayes classifiers, which are based on Bayes' rule and "naïvely" assumes independence—they are only valid to multiply probabilities when the events are independent (Witten and Frank, 2005).

The LB classifier was developed by Friedman and colleagues (2000) and performs an additive logistic regression and generates individual models in every interaction. The LogitBoost algorithm uses the cross-validation to determine an appropriate number of iterations and finish its execution before the performance comes down (Witten and Frank, 2005).

The AB1 classifier is an algorithm which is specifically designed to classify and which can be used also as a model in every classifying algorithm. This algorithm can handle weight instances, where the weight of an instance is a positive number. The weighted instances treatment changes the method of mistake calculation (Witten and Frank, 2005).

In addition to NBU, LB and AB1, the WEKA was used for data mining. WEKA was developed at the University of Waikato in New Zealand, and the name stands for *Waikato Environment for Knowledge Analysis*. This package has been implemented in the software language Java, and today stands out as probably the most competent and comprehensive package with algorithms of

machinery learning in academic and nonprofit world (Machine Learning Group, 2013). As a public data mining platform, WEKA collects many machine learning algorithms to mine data, including data pretreatment, classification, regression, cluster class, association rule mining and visualization on new interface. WEKA is very important for the academic and industrial researchers, and it could be used for teaching purposes.

Figures 2a and b depict the nominal and numeric variables of the newly graduated students here involved.

Selected attribute			
Name: **Gender**			Type: Nominal
Missing: 0 (0%)	Distinct: 2		Unique: 0 (0%)

No.	Label	Count	Weight
1	Female	107	107.0
2	Male	73	73.0

(a) Nominal Variable - Gender

Selected attribute			
Name: Age			Type: Numeric
Missing: 0 (0%)	Distinct: 9		Unique: 0 (0%)

Statistic	Value
Minimum	18
Maximum	27
Mean	23.761
StdDev	1.732

(b) Numeric Variable - Age

Fig. 2 Information displayed in the *Selected attributes* section

The cross validation on the classifier is very important for the evaluation of the classifier's robustness (Han *et al*., 2001; Kantardzic, 2011). In the present investigation, a 3-fold cross validation consisting of a randomly split data set into 3 subsets of equal size was used. Two subsets were used for training. One subset was used for cross validating, and one for measuring the predictive accuracy of the final constructed network. This procedure was performed three times so that each subset was tested once. Test results were averaged over 3-fold cross validation runs. Data splitting was carried out without sampling stratification. The WEKA software toolkit can calculate all these performance metrics after running a specified k-fold cross-validation. The prediction accuracy of the models was compared.

3. RESULTS AND DISCUSSIONS

Once the chosen algorithms were executed, the results were assessed and analyzed and the data are here reported clearly.

3.1 The input variable analysis

Graphs for each input variable in relation to the employed class are in the Figure 3 depicted. In these graphs, the dark part shows the level of the unsuccessful group (in this dataset it shows the unemployment), while the lighter part shows the successful group. In the present investigation 180 newly graduated students are involved. 100 students were from Tirana and the remainder from Korça. The following data could be obtained from the graph 3: i) Tirana marking the higher employment rate, ii) the young generation today prefers to attend the university. 17 out of 180 students dropped out of education level. 75 attend bachelor classes, 66 students attend master degree classes and 2 are PhD students. The 2 PhD students are already employed. In addition to PhD students, the master degree students are employed as well, iii) the employment-education relationship shows that most of the students that have finished the high school are employed, iv) most of the students attend the Albanian universities (155 students). However, the students studying abroad have more employment opportunities, v) the number of students attending private universities is higher (58). In addition, they have more employment opportunities. Grade average merely has any impact on employment opportunities, vi) few students foster intellectual pursuits (training, courses etc), vi) the more foreign languages, the more employment opportunities and, vii) employment rate is high (126 graduated students).
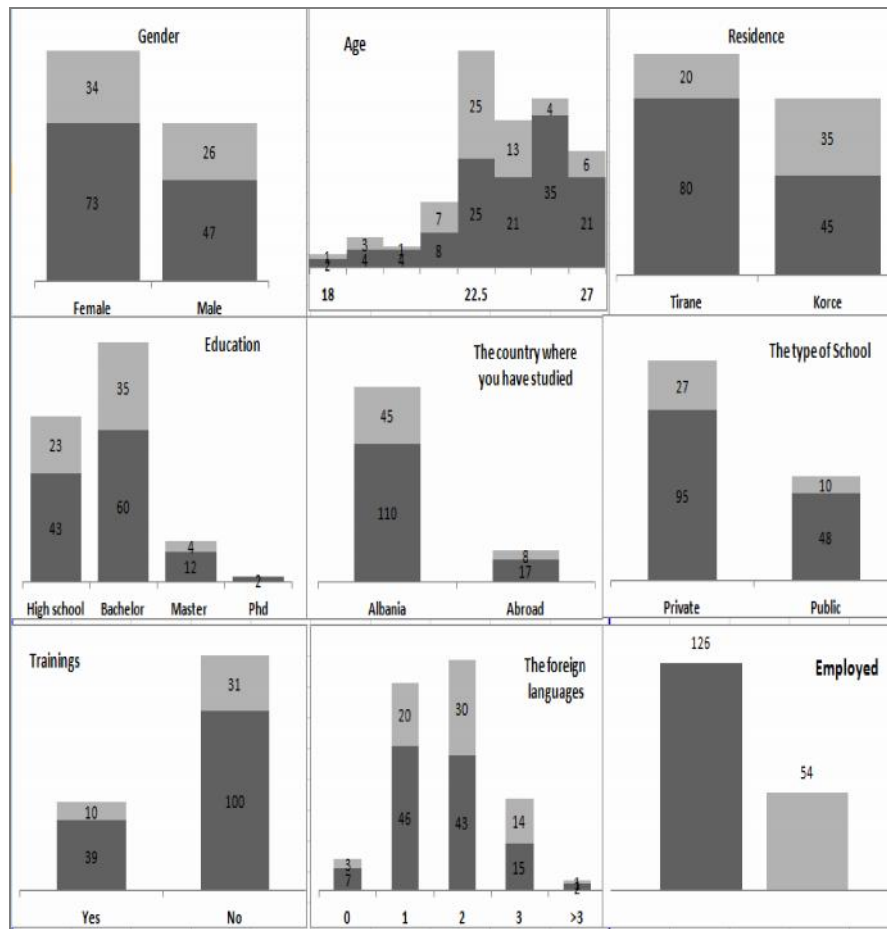
Fig. 3: Graphs for each input variable in ration with the employed class

3.2 Evaluation of the algorithms

In the present investigation, three data mining algorithms were involved to evaluate the performance and the efficiency of the algorithms here involved and the results are in the table 2 reported.

Table 3 reports the data obtained from the three algorithms   true positive (TP), true negative (TN), false positive (FP), false negative (FN), precision and the recall (Witten and Frank, 2005).

TP rate is the proportion of examples which were classified as class X among all examples which truly have class x. In the present investigation, it represents the ratio between the data predicted as positive and all the data included in the "Employed" class.

TN represents the ratio between the data predicted as negative and all the data included in the "Employed" class.

FP rate is the proportion of examples which were classified as class *x*, but belong to a different class, among all examples which are not of class *x. In the present investigation it* represents the ratio between the data classified by mistake in the "Employed" class and all the data included in all other classes, except the "Employed" class.

FN is the number of examples predicted negative that are actually positive

Precision is the proportion of positive predictions that are correct. *In the present investigation it re*presents the ratio between all the examples that indeed belong to the "Employed" class and the total number of the examples classified as belonging to the "Employed" class.

Recall is the proportion of positive labeled instances that were predicted as positive:

$$Recall = \frac{TP}{TP + FN}$$

F-measure a combined measure for precision and recall:

$$F - Measuare = \frac{2 * Precision * Recall}{Precision + Recall}$$

The performance of the three models was evaluated based on: i) prediction accuracy, ii) the time to build the model and, iii) the error rate (Table 2).

Table 2 The evaluation of the performance algorithms

Evaluation Criteria	Classifier		
	NBU	LB	AB1
Time taken to build model (Sec)	0	0.03	0.01
Correctly Classified Instances	133	132	133
Incorrectly Classified Instances	47	48	47
Accuracy of prediction	73.89%	73.33%	73.89%
Kappa Statistic	0.3143	0.3324	0.3614
Mean Absolute Error	0.3386	0.346	0.3576

Table 3 Detailed accuracy by Class

Classifier	TP	FP	Precision	Recall	F-Measure	Class
NBU	0,833	0,481	0,802	0,833	0,817	Yes
	0,519	0,167	0,571	0,519	0,544	No
LB	0,865	0,574	0,779	0,865	0,820	Yes
	0,426	0,135	0,575	0,426	0,489	No
AB1	0,865	0,556	0,784	0,865	0,823	Yes
	0,444	0,135	0,585	0,444	0,505	No

The Confusion Matrices are very effective in the evaluation of the classifiers. The columns stand for the prediction and the lines for the current classes. The confusion matrix on the Table 4-6 showed that NBU, LB and AB1 algorithms provided relatively good results. As a result, data mining can help in the prediction of the youth success in the job market (employed: YES or NO).

Table 4 NaiveBayesUpdatable Confusion Matrix

NBU		Predicted Class	
		A-Yes	B-No
Actual Class	A – Yes	105	21
	B- No	26	28

Table 5 LogitBoost Confusion Matrix

NBU		Predicted Class	
		A-Yes	B-No
Actual Class	A – Yes	109	17
	B- No	31	23

Table 6 AdaBoost1 Confusion Matrix

NBU		Predicted Class	
		A-Yes	B-No
Actual Class	A – Yes	109	17
	B- No	30	24

3.3 RESULTS

In the present paper, three important Data Mining algorithms were involved to evaluate the preprocessing of the data and their performance. In the end, and the most appropriate algorithm was determined based on the accuracy.

Results reported that: i) the LogitBoost classifier was the algorithm with the lowest accuracy (72%) and the longest execution time (0.03 sec), compared to

the other two algorithms. Consequently, it is inappropriate for application, ii) the AdaBoostM1 algorithm had the same accuracy (73.88%) as the NaiveBayesUpdatable, but its execution time was longer (0.01 sec). The AdaBoost1 algorithm managed to classify accurately more instances of the class "Yes" (109) comparing to the NaiveBayesUpdatable algorithm, which managed to accurately classify only 105 instances from the 126 examples belonging to the class "Yes", ii) based on the Table 2 and 3, the NaiveBayesUpdatable algorithm not only has a shorter execution time, but it offers higher accuracy, especially for the examples of the "No" class in which the other two algorithms failed. In addition, the average value of TP, Precision, Recall, the F-Measure and MCC were higher compared to the other cases, while the value of FP (which is not a positive indicator) is smaller compared to the classification obtained from the AdaBoost1 algorithm. Even though the algorithms have the same accuracy of 73.8889%, the data obtained from the output make the NaiveBayesUpdatable the most appropriate algorithm. In addition, it can simultaneously be correct and user-friendly for its users. Moreover, it could be used to: i) assess the employment opportunities, ii) the unemployment rate, iii) identify the job market requirements, etc. It could be used by the companies when hiring. The cause- consequence relationship is a means to address appropriate education policies.

4. CONCLUSIONS

Graduates employability remains a national issue due to the increasing number of graduates produced by higher education institutions each year. Consequently, the present investigation aims at: i) assisting education state institutions in encouraging students to foster their intellectual pursuits, ii) determine the factors effecting employment, iii) predicting whether a graduate has been employed or remains unemployed after graduation based on a series of classification experiments involving various algorithms under Bayesian and Meta methods to classify a graduate profile as employed or unemployed. Results showed that the NaiveBayesUpdatable algorithm, a variant of the Bayes algorithm, provided the highest accuracy -73.89%. The average accuracy of the two other algorithms was 73.61%.

The cause-consequence relationship is a means to address appropriate education policies based on job market tendencies.

As the data here reported are relevant for the state institutions involved in the area, the investigation must further by collecting a real and large data set from different sources such as graduate profiles from university student database and apply the model using such data. In addition, as the correlation factor between current and previous employers or the reason of choosing one

job to another will be included. Moreover, other classification methods could be applied to test the most suitable method for the structure of the graduate employability in Albania and give better classification accuracy.

REFERENCES

Al-Radaideh QA, Nagi EA. **2012** *Using data mining techniques to build a classification model for predicting employees performance*, International *Journal of Advanced Computer Science and Applications*, **3(2)**; 144.

Çela A, Kamberi G, Pici E (FES Coordinator). 2015. *Albanian Youth 2015 "Slow change, Internet dependency and EU trust"*. "Friedrich Ebert" Foundation and IDRA Research & Consulting.

Friedman J, Hastie T, Tibshirani R. 2000. *Additive logistic regression: a statistical view of boosting*. Annals of Statistics, **28(2);** 337–407.

Han J, Kamber M, Pien J. 2001. *Data mining Concepts and Techniques*, 3rd ed. Morgan Kaufmann, Academic Press.

Kantardzic M. 2011. *Review of Data Mining: Concepts, Models, Methods and Algorithms*, 2nd ed. Institute of Electrical and Electronics Engineers.

Muntean C, Moldovan D, Veres O. 2010. *A data mining method for accurate employment search on the web*. Communication and management in technological innovation and academic globalization.

Romero C, Ventura S. 2007. *Educational Data Mining: a Survey from 1995 to 2005*, Expert Systems with Applications, Elsevier, 135-146.

WEKA - *Machine Learning Group at University of Waikato*, 2013. Available on: http://www.cs.waikato.ac.nz/ml/weka

Witten IH, Frank E. 2005. *Data Mining: Practical machine learning tools and techniques*, 2nd ed. Morgan Kaufmann, San Francisco.

Zhang Y, Oussena S, Clark T, Hyeonsook K. 2010. *Use data mining to improve student retention in higher education – a case study*. International Conference on Enterprise Information Systems, 190-197.

APPLICATION OF DIFFERENTIAL INTEFEROMETRY FOR ANALYSIS OF GROUND MOVEMENTS IN ALBANIA

Neki FRASHËRI
Polytechnic University of Tirana, Albania
Salvatore BUSHATI and **Gudar BEQIRAJ**
Academy of Sciences of Albania
Mentor LAMAJ
Albanian Geological Service

ABSTRACT

Albania is experiencing some geomorphological changes of great impact for the environment due to anthropogenic activity. These changes are considerably reflected in the tectonic zone of Pre-Adriatic Depression. Here, the area is characterized by erosion, landslides and movement of the Adriatic Sea shoreline.
Anthropogenic activities have been intense in the hilly ranges running parallel to the sea shore which partly consist of lousy sandstone due to some catastrophic landslides occurred. Complicated geomorphologic phenomena could be noted in the sea shore which is characterised by significant sea transgression in typically accumulative areas. These phenomena have been investigated for a long time using satellite imagery involving free archives of LANDSAT [Kotor], MODIS [Balwois], SAR ERS and ENVISAT. In the present paper, some results recently obtained using SAR images are reported.

1. REVIEW OF GEOMORPHOLOGIC PHENOMENA IN ALBANIAN LOWLANDS

Half of Northern of the Western Albania– the Pre-Adriatic Depression tectonic zone–is characterized by lowlands and hilly ranges, bordered in the West by the Adriatic Sea (Geological Map 1:200000 of Albania; Aliaj *et. al.*, 2001; Favreto *et. al.*, 2013; Frashëri *et. al.*, 2014). Once a seaside, quaternary deposits consisting of lousy sandstone, conglomerates and clay are recent geologic formations are located today 10-15 km inshore. Drini River and Mati River flowing in North and Shkumbini, Semani and Vjosa River flowing in South have been of great impact for the development of lowlands.

In the area of economic activities occurred in the Adriatic shore areas, the last twenty years mark a significant increase of uncontrolled buildings' construction in the hilly ranges harming the environment. Consequently, sea transgression has occurred and landslide intensified (Aliaj *et. al.*, 2001; Frashëri *et. al.*, 2014).

False colour LANDSAT images in the figure 1 represent respectively the southern (a) and the northern (b) sections of Albanian lowlands. Images are obtained by the combination of NIR bands from three different years 1972-1986-2002 as a single RGB image, where sea regression and transgression areas are visible in Blue and Red colours, respectively (Colesanti *et. al.*, 2003(a); Mannunta *et. al.*, 2003; Catani *et. al.*, 2006; Farina *et. al.*, 2006; ESA 2000-2015(a); ESA 2000-2015(b); Tamburini *et. al.*, 2011)
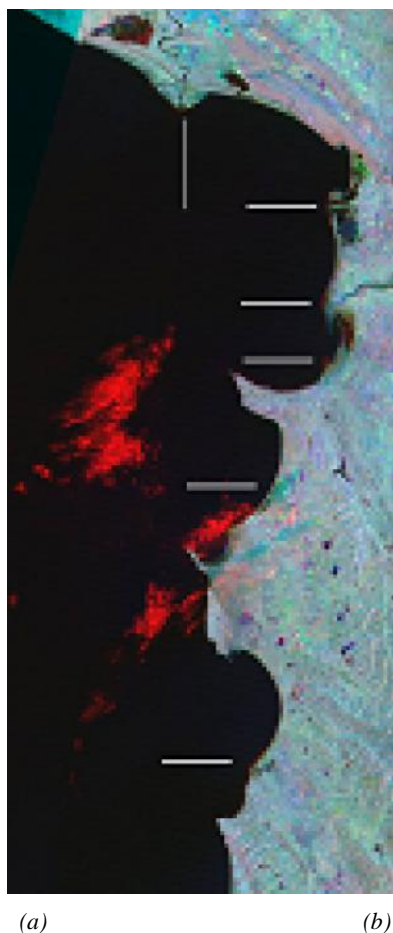

(a)   *(b)*

**Fig.1**. Lowlands from NIR bands of LANDSAT.

The Investigation was carried out in two "layers" in the following areas: i) the delta of Shkumbini River and the Durrësi region where the SqueeSAR™ software was used and multimage SAR calculations were carried out by the company TRE in Milano, Italy. Traditional two-image interferograms were calculated for both segments of lowlands as presented in figure 1 and, ii) Synej.

Figure 2 depicts a typical landslide which occurred in 2007 in Synej. These hills geologically consist of lousy sandstone. Anthropogenic activities have been of great impact for their structure. The sliding mass was separated from the hill forming a deep "canyon" that now is used as a road.

Fig. 2: Landslide in Synej.

Figure 3 depicts the section of the Adriatic beach in Semani delta. Here, water tower now situated offshore could be noted. The source of the sea transgression is the sea erosion and ground subsidence.
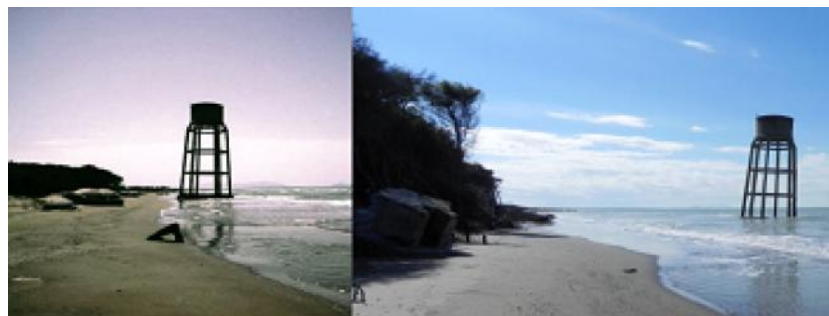
(a - 2007)   (b - 2010)

Fig. 3: Sea transgression in Semani area.

Once traditional images from the field and satellites LANDSAT AND MODIS were applied, SAR images and interferometry techniques were used to evaluate vertical movements of the ground with centimetric precision (Avallone and Zollo 1999; Osmanoglu, 2010; UGS, 2010; Moisiu *et al*., 2014). The methods applied are in the forthcoming section reported.

2. THE SAR METHOD

The SAR differential interferometry involving the surface interferometry and scatterer-based point-wise interferometry are in the present investigation used. The EOLISA software was used to access the database and select pair of images for both southern and northern segments of lowlands.

This method was applied based on (Ferreti *et al.*, 2001; Allievi et al., 2003 (a) Allievi *et. al*., 2003(b); Colesanti *et al*., 2003; Ubisci *et al*., 2005; Bottero *et al*., 2006; Albaviator, 2007; Reigber *et al*., 2007; Ruiz *et al*., 2008; Ferreti *et al*., 2011; Favretto *et al*., 2013; Giannico 2013) .

For surface interferograms formally suitable pairs of ASAR images were obtained from the ESA in the framework of a small project "Seman 14921", in 2013, to select pairs with the largest time difference possible. For the southern segment two pairs of IMS images were selected for the dates 2003/07/04: 2007/02/23 with baseline 40m and 2003/03/21: 2007/11/30 with baseline 746m. For the northern segment only one pair of IMS images was selected for the dates 2003/03/21: 2004/11/05 with baseline 100m. In addition, ERS images with larger time difference were downloaded. The problems with the orbit files prevented us from their use.

The calculation procedure was tested using the ASAR IMS images from L'Aquila earthquake of 2009. Two software packages were tested, firstly RAT & IDIOT and finally ESA NEST.

Two final interferograms were calculated using ESA NEST software with pairs of dates 2003/07/04: 2007/02/23 and 2003/03/21: 2007/02/23. The image of 2007/02/23 was not used because of related orbit file problem. Both wrapped and unwrapped interferograms were calculated. Interpretation of results is presented in the next section.

Point-wise scattered interferograms were calculated with the assistance of the T.R.E co., Milano, Italy. Here, the software package SqueeSAR was used (Fig. 4).
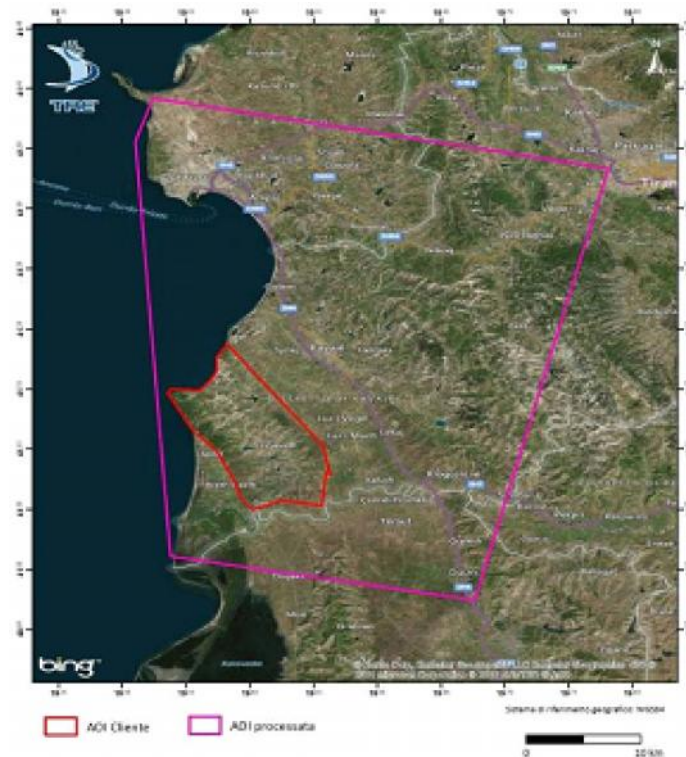
Fig. 4. The investigated area via SqueeSAR.

Two sets of ERS images were used for the period running from 1992 to 2000: i) 31 images of upward pass and, ii) 50 images of download pass. Time interval between two successive images was 35 days. A total of 11.657 scattered points was identified. Both horizontal and vertical velocities of scattered points were calculated from the Line-of-Soght (LOS) relative displacements.

3. RESULTS OF POINT INTERFEROGRAMS

Figure 5 depicts the points selected from the area covered by SqueeSAR for detailed analysis.

The most critical section of this area is located in the hilly range bordered with red at the centre of the image. Figure 6 depicts the numerous landslides identified in this section using traditional geological field work.
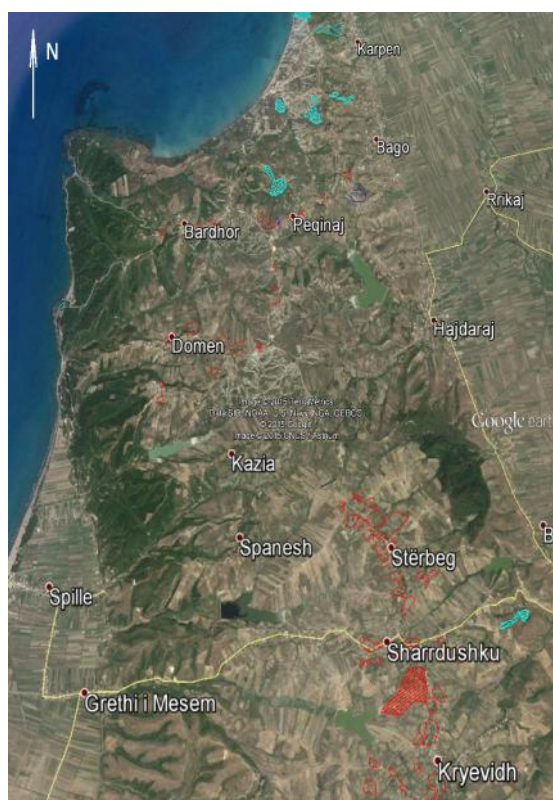
Fig. 5. Analysed areas from SqueeSAR

Fig. 6. Landslides identified through field work.

Annual average values of displacement velocities in LOS are in the figure 7 depicted.
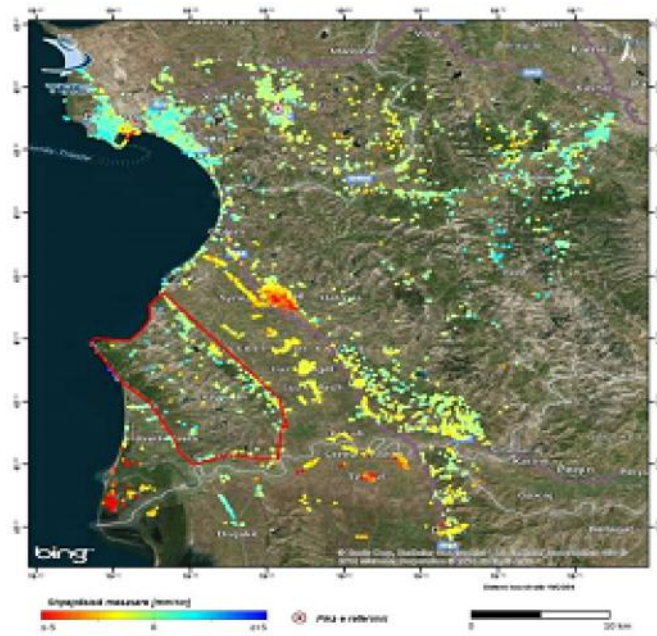

Fig.7: Annual averages of velocities in LOS.

Identified scattered points resulted mainly in urban areas along with those with scattered vegetation. Significant displacements are in Kavaja city, delta of Shkumbini River and the village of Tërbufi. Based on LOS displacements and orbital parameters the vertical and horizontal displacement velocities were calculated (Fig. 8 and 9).
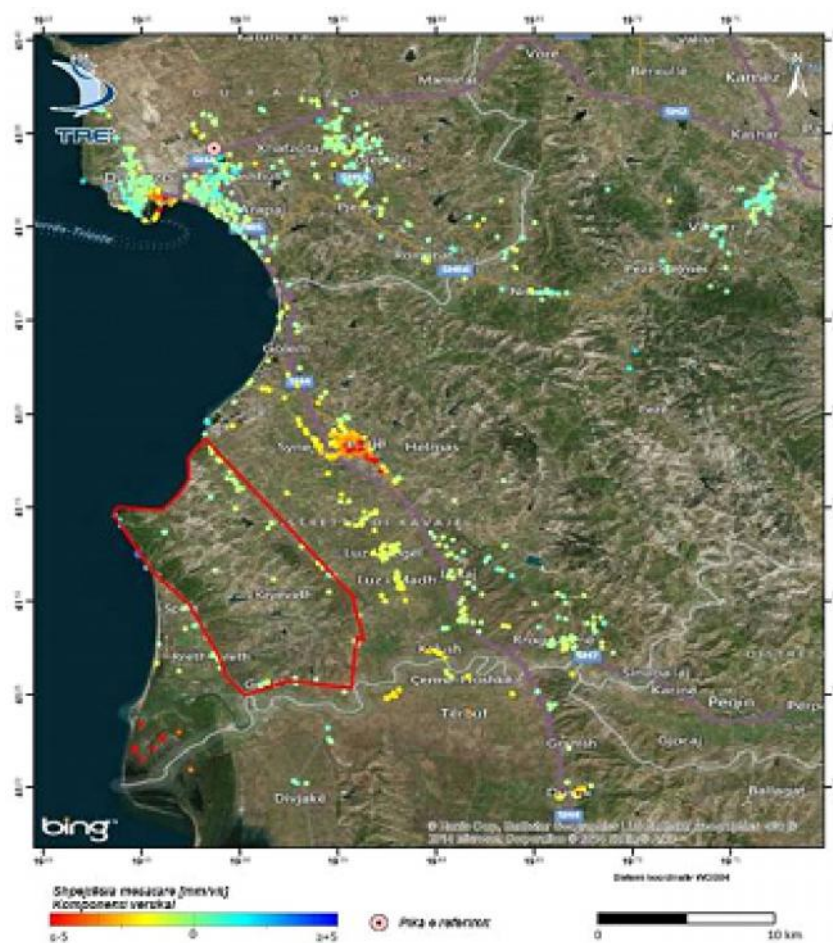
Neki FRASHËRI, Salvatore BUSHATI, Gudar BEQIRAJ,
Mentor LAMAJ

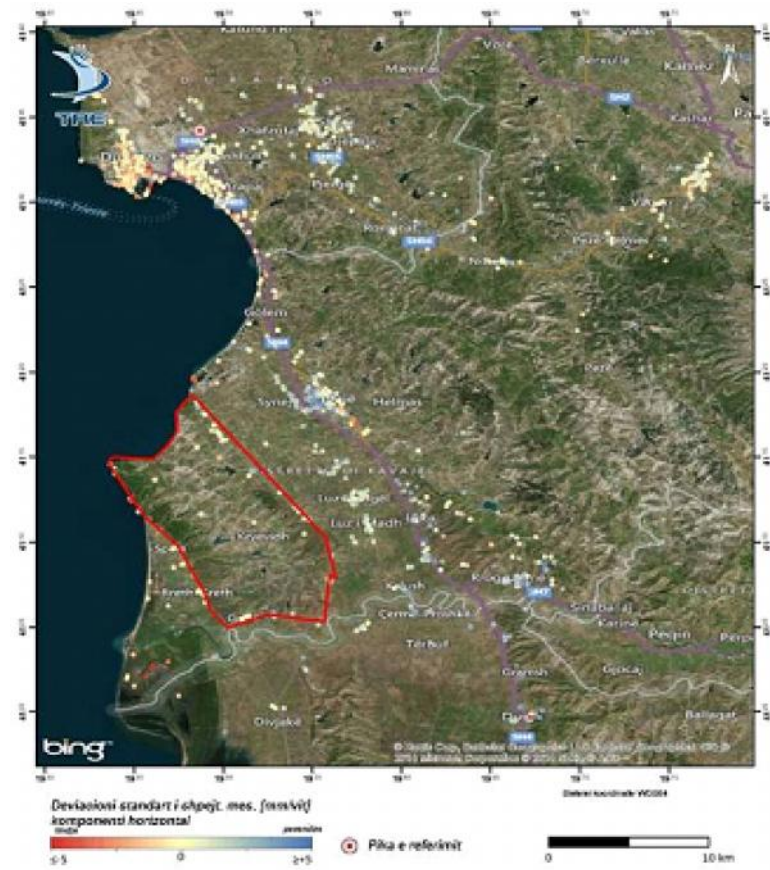**Fig. 8**. Annual averages of vertical velocities

Fig. 9: Annual average of horizontal E-W velocities.

The area bordered with red line has been subject to intense field work for identification of landslides and field results were compared with the SqueeSAR data. Satellite data helped identify new landslide sites in villages Sterbeg (Y – 43 74 489, X – 45 54 987; Fig. 10), Kazie-1 (Y – 43 73 673, X – 45 55 103; Fig.11) and Kazie-2 (Y – 43 73 617, X – 45 55 199; Fig. 12).
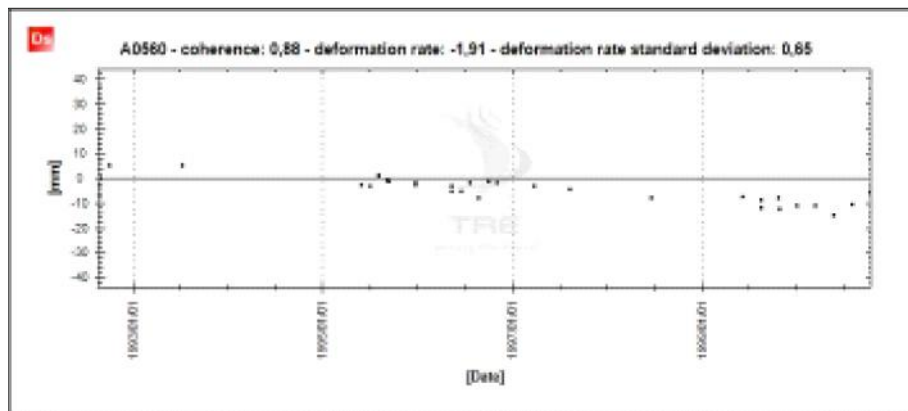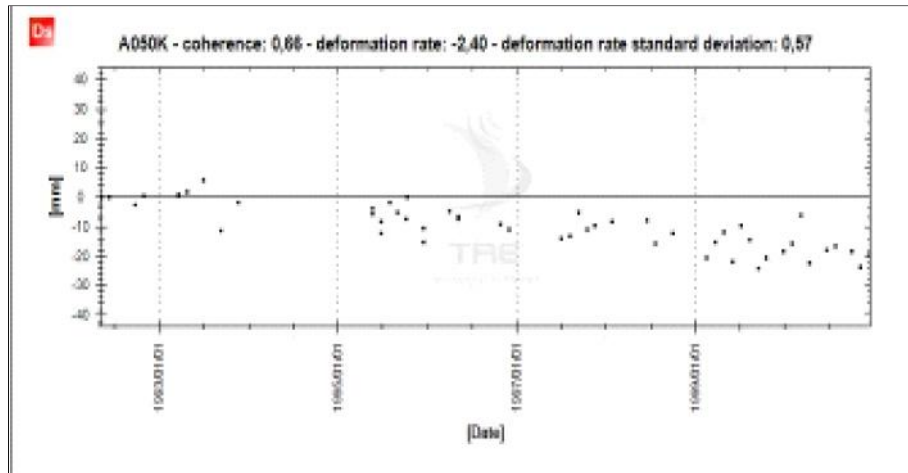
Fig. 10: Landslide in Sterbeg.

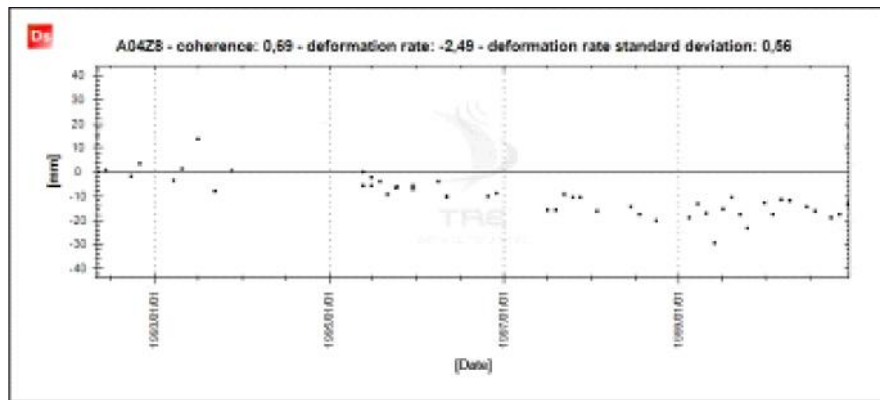**Fig. 11:** Landslide in Kazie-1.

Fig. 12: Landslide in Kazie-2.

4. RESULTS OF SURFACE INTERFEROGRAMS

The Figure 13 depicts the first differential interferograms calculated for the southern segment of the PreAdriatic Depression lowlands. The coherence is in the Figure 14 depicted. Few fringes which cannot be directly interpreted as vertical displacements of the ground could be met. For most of them the source is the vegetation variations, mostly in mountainous areas. Only fringes in the seashore in the point B and C are formally correlated with sea regression and transgression. The point A represents a wetland near the seashore. In Polis mountain range vegetation, variation is already identified via Landsat images.
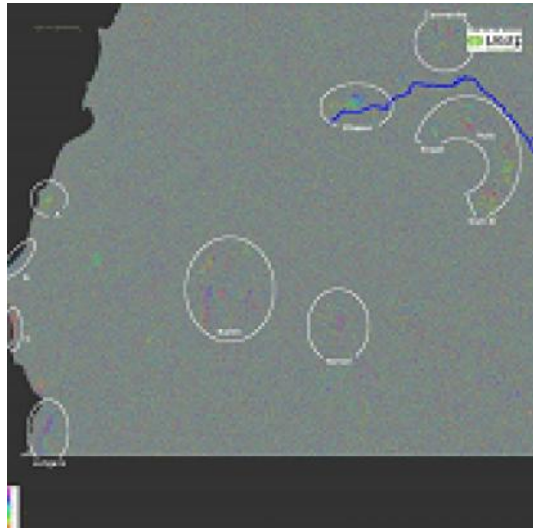
Fig. 13: Southern differential interferogram.
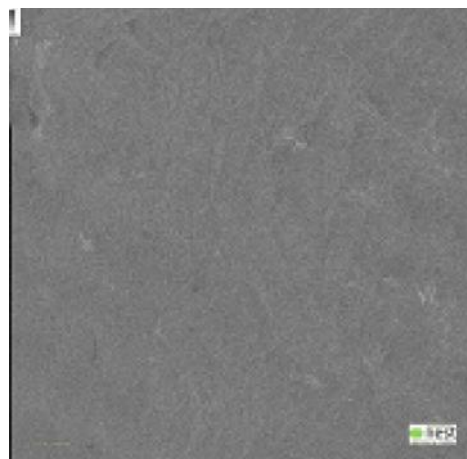
Fig.14. Southern interferogram coherence.

The unwrapped interferogram is in the figure 15 depicted. The image was filtered with Gaussian filter and a number of straight features correlating with the principal probably tectonic mountain ranges, fractures were identified.

Fig. 15. Unwrapped filtered interferogram.

The differential interferogram of northern segment of Pre-Adriatic Depression lowlands is in figure 16 depicted.

Fig. 16. Northern differential interferogram

Correlation of fringes with mountain ranges is more visible than in the south. Here, their interpretation relates to vegetation variations. Nevertheless, weak signs of correlation with the slow but huge landslide of Rragami in shores of Vau i Dejës Hydropower Lake are identified (Fig. 17).

(a)   *(b)*

Fig. 17: DEM (a) and interferogram (b) of Ragami landslide.

5. CONCLUSIONS

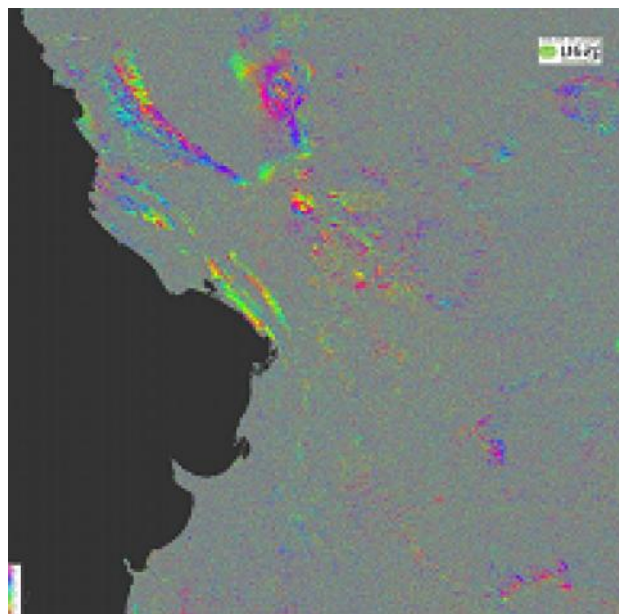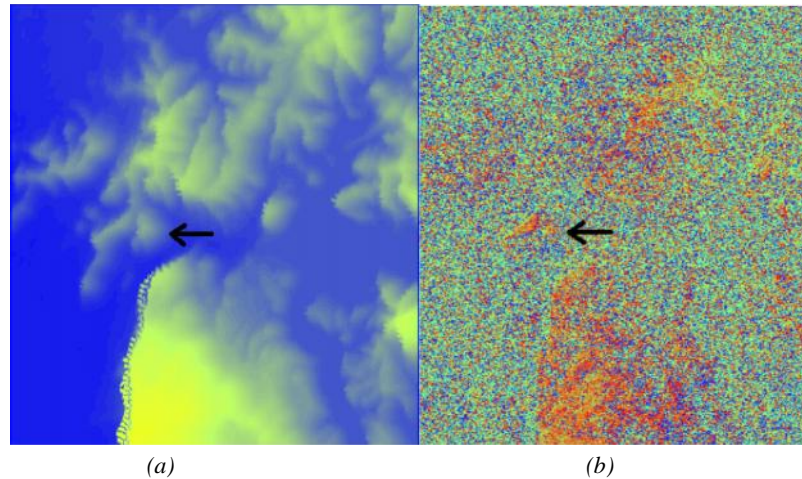Field observations in the PreAdriatic Depression lowlands reported that residential buildings are constructed in the Semani area, in a dynamic terrain being developed. Consequently, the buildings are partly destroyed as result of a complex of factors that include movements of deltas, possible subsidence of the ground evolved in several decades, and disastrous landslides.

There is a lack of scientific information regarding local ecosystem processes and inappropriate policies have been a continuous obstacle for suitable environmental solutions (Moisiu et al., 2014).

LANDSAT images helped identify the areas where the sea shore line has moved mainly inland, but the images resolution is inadequate for a detailed analysis of the shore and landslides in hilly areas. Radar imagery ERS and ENVISAT have resolution similar to the LANDSAT. They help evaluate the surface and point-wise of vertical movement of the ground in areas undisturbed by vegetation variations.

SqueeSARTTM and NEST software was used to develop interferograms. Multi-image pointwise interferograms built by the T.R.E. co. involving the SqueeSAR TM for a limited segment of lowlands allowed the identification of ground displacements related to landslides including unknown landslides in a studied area. The latter has been carefully scanned during the field works.

Subsidence was identified in some of the urban areas and in the delta of Shkumbini River in particular. Periodic use of point-wise interferograms for the entire territory would help to identify unknown, existing and potential landslides. In addition, appropriate measures for mitigation of negative impact could be taken.

A minimal number of SAR images provided by the NEST software were used for surface differential interferograms, which provided few fringes in hilly and mountainous ranges along with potential fringes to be caused by vegetation variations. Only few weak fringes potentially due to beaches vertical movements and massive but slow landslide (in one case) were identified.

The conclusion is simple – despite the complexity of terrain evolution in Albania including rapid urban expansion and significant vegetation variations, SAR interferograms proved to be a tool for analyzing the existing and potential landslides, as well as areas with significant vertical movements of the ground. Application of Remote Sensing technique involving the SAR intereferograms for the entire territory of Albania is a means to address an accurate management of catastrophic phenomena.

REFERENCES:

Albaviator. 2007. Seman (photo). https://www.flickr.com/photos/albanianairforce/869073392.

Aliaj Sh, Baldassare G, Shkupi D. 2001. Quaternary subsidence zones in Albania: some case studies. *Bulletin of Engineering Geology Environment* **59**:313-318 Springer Verlag.

Allievi J, Bonsignore F, Cespa S, Colesanti C, Colombo D, Ferretti A, Morelli M, Pistocchi M. 2003. Analisi di fenomeni di deformazione superficiale sul territorio dell'autorità dei Bacini Romagnoli a partire da serie di dati radar satellitari elaborati con la tecnica dei diffusori permanenti. Atti della 7^ Conferenza Nazionale ASITA, Verona, 28-31 ottobre, 45-50.

Allievi J, Ambrosi C, Ceriani M, Colesanti C, Crosta GB, Ferretti A, Fossati D. 2003. Monitoring slow mass movements with the permanent scattered technique. Geoscience and Remote Sensing Symposium, IGARSS '03. *Proceedings of 2003 IEEE International*, Volume: **1 215-217**.

Avallone A, Zollo A. 1999. Subsidence of Campi Flegrei (Italy) detected by SAR. *Geophysical Research Letters,* 26(15).

Bottero D, Poggi F, Cespa S. 2006. Il monitoraggio SATELLITARE. Primi risultati in Liguria della sperimentazione dell'utilizzo della tecnica PSInSARTM. Quarry & Construction, Ottobre 2006.

Catani F, Colombo D, Farina P, Fumagalli A, Kukavicic M, Marks F,

Menduni G, Moretti S. 2006. Utilizzo di dati telerilevati nella mappatura e nel monitoraggio dei fenomeni franosi e nell'analisi della suscettibilità da frana. *Giornale di Geologia Applicata.* **3**; 173-180.

Colesanti C, Ferretti A, Prati C, Perissin D, Rocca F. 2003. ERS-ENVISAT permanent scatterers interferometry. Geoscience and Remote Sensing Symposium, IGARSS'03. *Proceedings IEEE International*, **2**, 21-25.1130 - 1132 July.

Colesanti C, Ferretti A, Prati C, Rocca F. 2003. Monitoring landslides and tectonic motions with the Permanent Scatterers Technique. 2002., *Engineering Geology,* Elsevier.**68**, 3–14.

ESA Earth Online 2000 – 2015. EOLi – ESA's Link to Earth Observation. https://earth.esa.int/web/guest/eoli.

ESA 2000 – 2015. NEST - Next ESA SAR Toolbox. https://earth.esa.int/web/nest/home.

Farina P, Colombo D, Fumagalli A, Marks F, Moretti S. 2006. Permanent Scatterers for landslide investigations: outcomes from the ESA-SLAM project. Elsevier, *Engineering Geology,* **88**; 3–4.

Favretto A, Geletti R, Civile D. 2013. Remote Sensing as a Preliminary Analysis for the Detection of Active Tectonic Structures: an Application to the Albanian Orogenic System. GEOADRIA 18/2, 97-111.

Ferretti A, Fumagalli A, Novali F, Prati C, Rocca F, Rucci A. 2011. A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing*, **49 (9);** 1-11.

Ferretti A, Prati C, Rocca F. 2001. Permanent Scatterers in SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, **39 (1);** 8-20.

Frashëri N, Beqiraj G, Bushati S, Frashëri A, Taushani E. 2014. Remote sensing analysis of the Adriatic shoreline movements. International Scientific Conference on Integrated Coastal Zone Management in the Adriatic Sea. Institute of Marine Biology, Kotor, Montenegro. 29 September – 1 October 2014.

Giannico C. 2013. Il monitoraggio satellitare: nuova frontiera per la studio di strutture e infrastrutture. 2013. Il Giornale dell'Ingegnere, N. 10 - Ottobre.

Manunta P, Brugioni M, Casagli N, Colombo D, Deflorio AM, Farina P, Ferretti A, Gontier E, Graf K, Haeberle J, Lateltin O, Meloni E, Mayoraz R, Montini G, Moretti S, Paganini M, Palazzo F, Spina D, Sulli L, Strozzi T. 2003. SLAM, a service for landslide monitoring based on EO-data. *Proceedings of FRINGE.* Workshop, Frascati, Italy, 1 – 5 December (ESA SP-550, June 2004).

Moisiu L, Kapsimalis V, Durmishi Ç, Anagnostou C. 2014. Application of DPSIR Framework to a Mediterranean Transitional Water Body: The Butrint

Lagoon, SW Albania. *Journal of Natural and Technical Sciences. (JNTS)*, Academy of Sciences of Albania. **1(3)** XIX; 96-112.

Osmanoglu B. 2010. Tutorial: Processing envisat SAR interferograms with ADORE. http://www.osmanoglu.org/adore-doris-downloads.

Reigber A, Erten E, Guillaso S, Hellw O. 2007. IDIOT: a free and easy-to-use software tool for dinsar analysis. *Proceedings of Envisat Symposium*. Montreux, Switzerland 23–27 April.

Ruiz AM, Sousa J, Hanssen RF, Perski Z, Gil AJ, Bastos L, Galindo-Zaldívar J. 2008. Measurement of ground subsidence in the Granada area (Southern Spain) using PS-InSAR. IX International Congress on Geomatics and Surveying Engineering, Valencia.

Tamburini A, Del Conte S, Larini G, Lopardo L, Malaguti C, Vescovi P. 2011. Application of SqueeSARTM to the characterization of deep seated gravitational slope deformations: the Berceto case study (Parma, Italy*). Proceedings of the Second World Landslide Forum*, 3-7 October, Rome.

Urbisci S, Allievi J, Cespa S, Gervasi C. 2005. Elaborazione di dati radar satellitari per la Carta del Rischio del Patrimonio Culturale della Lombardia. ASITA 2005, 15-18 Novembre, Catania, Italia; Pagine 1-6.

USGS (Sept 30, 2010) Shuttle Radar Topography Mission. http://srtm.usgs.gov/index.php.

USGS (12/30/2014) Landsat Data Access. http://landsat.usgs.gov/Landsat_Search_and_Download.php.

APPLICATION OF REMOTE SENSING FOR THE ANALYSIS OF ENVIRONMENTAL CHANGES IN ALBANIA

Neki FRASHËRI, Gudar BEQIRAJ and **Salvatore BUSHATI**
Academy of Sciences of Albania
Alfred FRASHËRI
Faculty of Geology and Mining, Polytechnic University of Tirana, Albania

ABSTRACT

The present paper reviews some of the remote sensing studies carried out to investigation environmental changes in Albania using simple methods and image processing software for general use and exploiting free internet archives of satellite imagery. Data of significant importance were obtained about the hot areas of environmental changes such as sea coasts experiencing sea transgression, temporal variations of vegetation and aerosols, lakes, landslides and regional tectonics.

1. INTRODUCTION

Satellite imagery is widely used to investigate and map environmental problems and changes, reducing at the same time the expensive field work.

Remote sensing is a means to address a wide range of environmental, climatic, atmospheric and oceanography studies. In the present paper, information about some environmental hot spots in Albania regarding variations of water bodies and sea shorelines, evolution of vegetation in years due to anthropogenic activity is reported and global situation of aerosols in time, delineation of regional tectonics related with destructive earthquakes etc., and in particular lake areas and Adriatic shore line is evaluated.

2. MATERIALS AND METHODS

LANDSAT and MODIS images were obtained from the online archive of Geological Service of the United States. Some experiments were made using SAR interferometry involving ENVISAT ASAR images obtained from the European Space Agency online archive. The methods (Frashëri *et al*., 2010) here used were: i) combination of different bands to identify specific

environmental features (combining of a specific band from 2-3 dates in a false RGB color image, and using color differences to identify temporal change) and, ii) comparison of images from different dates for a single area (calculation, for a temporal suite of images of a specific band, of linear or quadratic trend in time for each pixel, obtaining gray scale images representing the spatial distribution of average values of the specific band, and of respective spatial distribution of temporal gradients of each pixel).

Images processing involved professional software for remote sensing to cut images for selected areas, image processing of general use for combination of RGB channels, and in-house developed software for trend analysis in time domain for a suite of images of the same area.

Ohrid – Prespa Lakes area
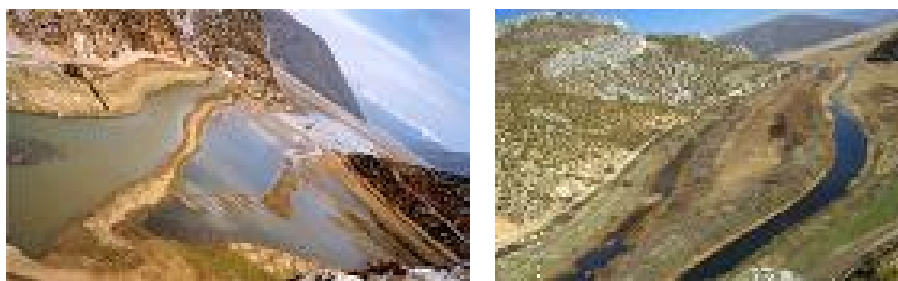
Fig.1: SW of Micro Prespa Lake in 2000 and 2006.

Ohrid – Prespa Lakes represent a single surface and underground water body system, with water flow from Micro Prespa into Macro Prespa, and from the latter into Ohrid Lake in Saint Naum and Tushemisht springs. Prespa Lakes have experienced significant changes (Frashëri *et al*., 2010; Frashëri et al., 2012; Frashëri *et al*., 2013). In the south-west, the Micro Prespa Lake is filled with sediments from Devolli River, due to artificial channels used for water management (Fig.1). The water surface of the Macro Prespa Lake was reduced several meters of height. Apparently, the source is the underground karst changes. In addition to field observations, satellite imagery was used to evaluate regional environmental changes from 1972 to present.

Fig.2: Ohrid-Prespa Lakes: Landsat GRNir band combination for years 1972 and 2001.
White circles show areas with loss of free water surface and development of reeds.

The combination of Landsat bands helped identify the dynamics of shorelines due to water loss in Macro Prespa Lake and sedimentation in Micro Prespa Lake. In the Figure 2 and 3, the combinations of Green-Red-NIr bands for 1972 and 2001 show clearly the loss of free water surface in Micro Prespa Lake. Combination as RGB of NIr bands from 1972, 1986, 2002, 2010 is depicted in Figure 4, where the dominance of Red indicates areas with loss of free water surface and a slight improvement in Micro Prespa Lake in 2010 when the use of artificial channel ended.
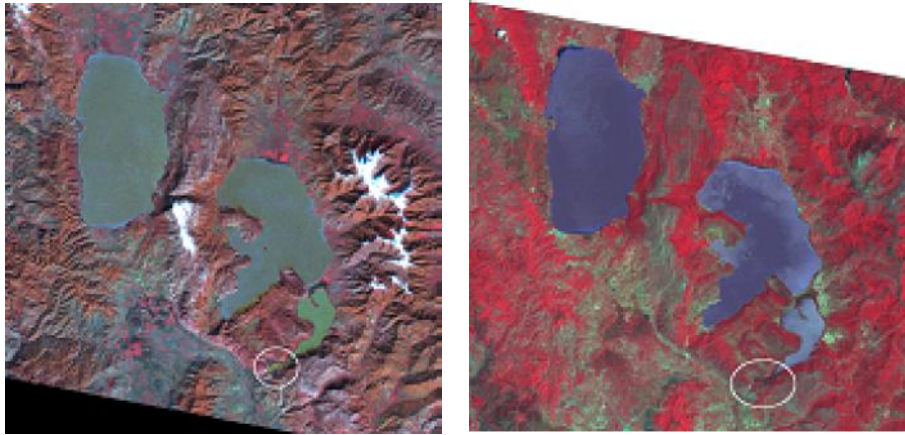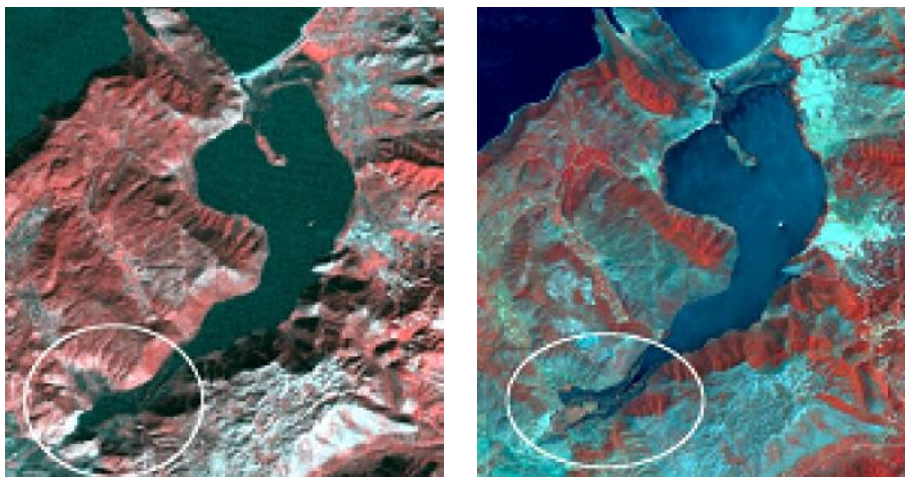

Fig.3: Micro Prespa Lake: Landsat BGNir band combination for years 1972 and 2001.
White circles show areas with loss of free water surface and development of reeds.
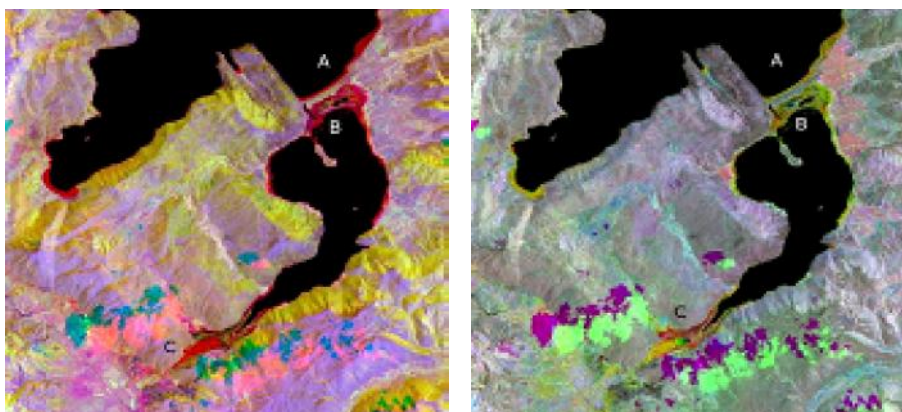
Fig.4: Micro Prespa Lake evolution from Landsat NIr images. Left: Blue~1972,Green~1986,Red~2002. Right: Blue~1986, Green~2002, Red~2010.

The trend analysis in time domain for a suite NDVI views provided by the Landsat images of the period 1972 – 2012 for Prespa lakes area is in Figure 5 depicted. Results gave trend coefficients for each pixel, presented with false colors. The image processing software of general use was involved for the classification of the gradient of change in order to identify areas where vegetation is improving.
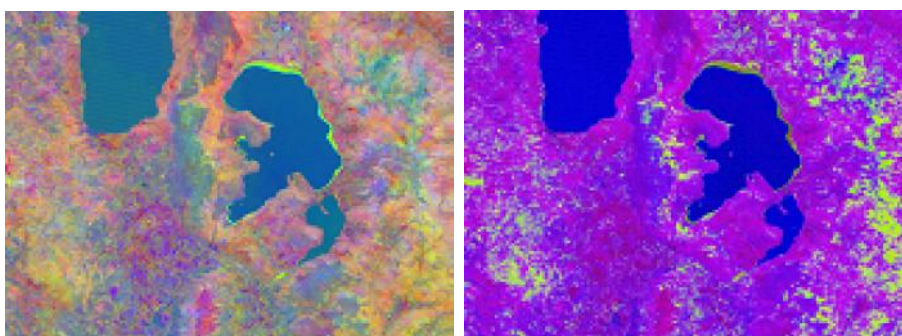
Fig.5: Trend analysis of NDVI for Landsat images of 1972-2012. Left: average (Red) and gradient (Blue-Green). Right: gradient enhanced, positive values in Blue-Green.

The Shkodra Lake area

Shkodra Lake is located in northwestern Albania and is the largest lake in the entire Balkan Peninsula, with an area of 368 hectares. The lake is shared between Albania and Montenegro, of which 149 hectares and 57 km of shoreline fall within the Albanian territory. The lake is connected with two important rivers – Drini and Buna. Buna River flows from the Lake into Adriatic Sea.

Drini River is the source of the complicated hydrological system which characterizes the area because it has shifted from Shëngjini Bay to Buna River, due to human intervention.
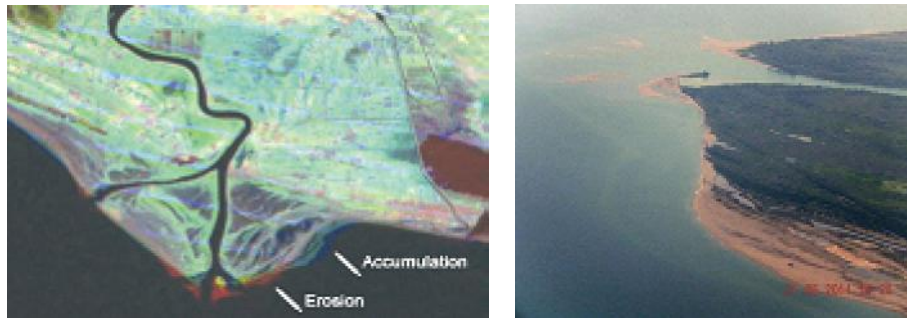
Fig.6: Left: changes of Adriatic shoreline in the delta of Buna River, combination of Landsat NIr bands for years 1975 (Red) and 2009 (Blue-Green). Right: aerial view of Buna River delta.

Figure 6 depicts the delineation of the erosion of delta of Buna River and with the accumulations in Velipoja beach obtained due to the combinations of Landsat NIr images from 1975 and 2009 (Frashëri *et al*., 2011; Frashëri *et al*., 2014).

The same method was used for the Shkodra Lake, delineating sedimentation areas in northern shore and near the area where Buna River flows. The source of the latter might be the Drini River which in winter flows into the Lake. Figure 7 depicts the Northern shores accumulation characterized by fast development of reeds beds.
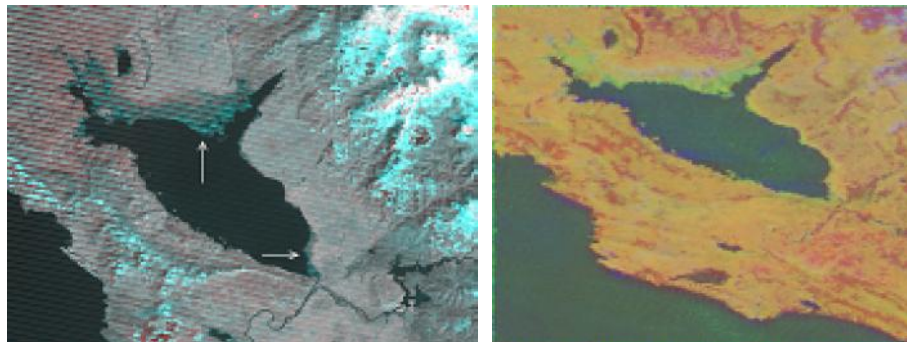
Fig.7: Left: sedimentation in Shkodra Lake – Landsat NIr bands combiantions for 1975 (Red), 2000 (Green) and 2009 (Blue). Right: NDVI trend analysis from Landsat images, averages as Red, gradients as Blue-Green (visible the increase of vegetation in new sediments in northern shores of Lake).

The Drini – Buna Rivers are mentioned for periodic floods in winter. Combination of NIr bands from MODIS was used to delineate flooded areas from these rivers in 2010 and recent floods from Vjosa River in the south (Fig.8)
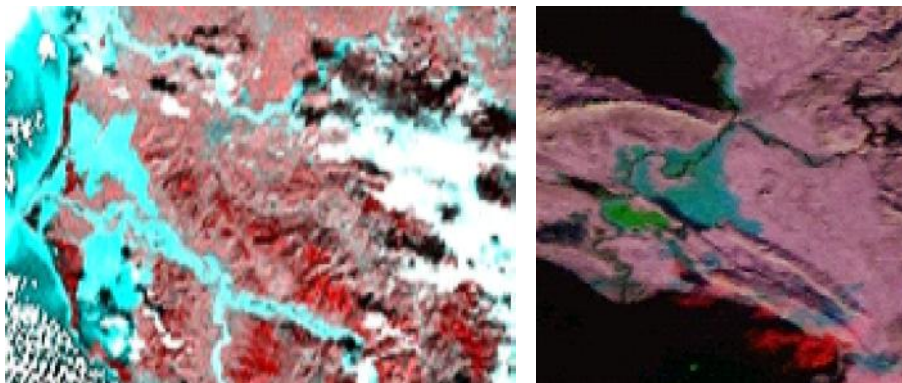
Fig.8: Delineation of flow areas through combination of MODIS NIr bands. Left: Vjosa River, Dec 2014 as Red, Jan 2015 as Blue-Green. Right: Buna River, Dec 2009 as Red, Jan 2010 as Blue-Green

Adriatic Sea coastal area

The Adriatic Sea coastal areas are mostly Quaternary accumulations, part of Preadriatic Depression tectonic zone.
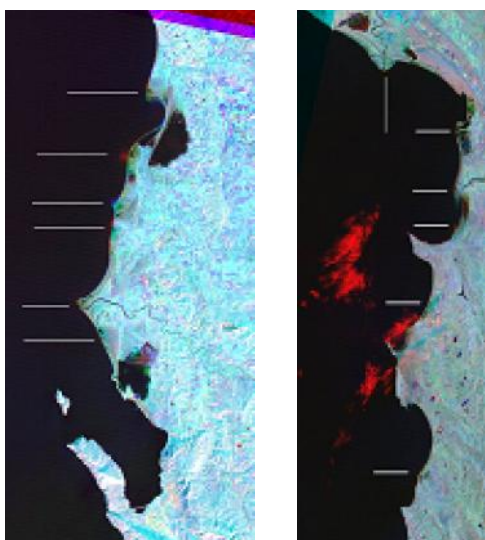
Fig.9: Landsat NIr band combinations for southern (left) and northern (right) parts of Adriatic coastal zone: 1972as Red, 1986 as Green and 2002 as Blue. White linesh point coastal sites where sea regression or transgression is observed.

Adriatic coast represents about two-thirds of country's sea coasts and is experiencing important changes related with regression and transgression of the sea, result of a combination of subsidence and erosion-accumulation that have strongly impacted human activities in respective areas (Frashëri *et al*., 2012; Frashëri *et al*., 2014; Frashëri *et al*., 2015). Comparison of Nir bands of Landsat images from 1972 were a means to address the delineation of regression and transgression areas, in particular the movement of rivers deltas (Fig.9).

Aerosols

Figure 10 depicts the global temporal variation of aerosols investigated using the trend analysis of a suite of MODIS images preprocessed for the optical depth of aerosols (Pano *et al.,* 2004).
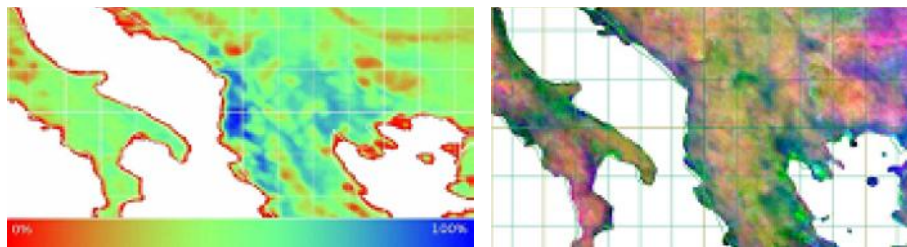
Fig.10: Coverage (left) and trend analysis (right: averages as Red, gradient as Blue-Green) of aerosols optical depth from MODIS images for the period 2000-2010.

Detailed analysis identified coastal areas with higher levels of aerosols (Fig.11 left), as well as some hot spots with an increasing gradient (Fig.11 right).
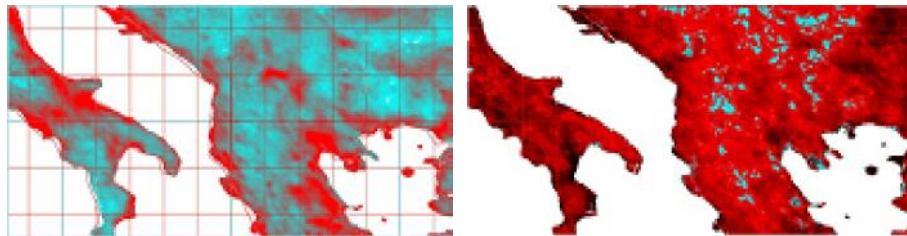
Fig.11: Trend analysis of aerosols optical depth. Left: the average (low values in Red, high values in Blue-Green). Right: the gradient (negative in Red, positive in Blue-Green).

The main source is the reduction of the total of aerosols for the whole column of atmosphere in most of the territory of the country. This does not mean a-priory that distribution of aerosols near the ground surface follows the same trends.

Vegetation

Figure 12 depicts the evolution of vegetation investigated via MODIS images involving the trend analysis of Normalized Difference Vegetation Index (NDVI) for the entire territory of the country (Frashëri *et al.*, 2010).
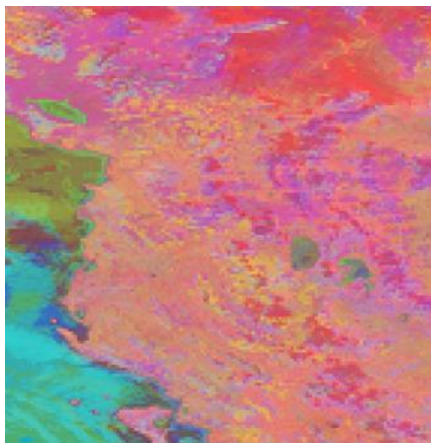
Fig.12: Trend analysis of NDVI from MODIS images (Red and NIr bands), period 2005-2010. Averages in Red, gradient in Blue-Green.

As it is clearly shown, there are complicated patterns of vegetation, especially in mountainous areas. In general an increase of vegetation index is observed, related with the abandonment of mountainous areas and increase of raining in summers. Analysis of forestry areas is not included in this study.

Application of Radar Interferometry

ESA Envisat ASAR images were used for the calculation of differential interferograms for the North-Western Albania. Clear fringes which delineation was correlated with near coast hilly ranges in Shkodra area were obtained (Fig.13). The source is the vertical movement of the ground and vegetation variations. The type of fringes delineation is related to the vegetation (Pano *et al.*, 2014).

(a)   (b)
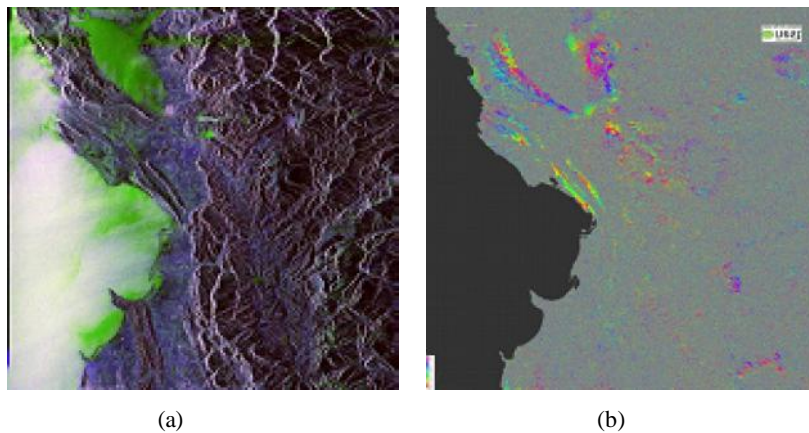
Fig. 13: Environmental changes in northern Preadriatic Depression from Envisat SAR imagery. Left: Combination of intensity bands from 2003/03/21 and 2004/11/05. Rigt: DINSAR inerferogram from the same images.

Geological Applications

Efforts for geological applications of remote sensing were made while investigating the Shkodra–Peja traversal fracture (Pano *et al.*, 2006).

Fig.**14**: Combination of Landsat bands: 7 as Red, 4 as Green, 2 as Blue. Red line represent the Shkodra-Peja fracture. Mirdita massif (magenta area in center) distinguishes from limestone of Alps north of the fracture Green areas represent the vegetation coverage.

A combination of Landsat bands 7 as Red, 4 as Green, 2 as Blue was used to detect for the changes, soil types and moisture (Fig.14). Tectonic zone of ophiolites of Mirdita is clearly separated from limestone Alps. The separation line coincides with the fracture as determined from other studies.

Geophysical investigation was carried out for slope stability and landslides occurring in Ragami in the hilly shores of Vau Dejes artificial lake involving the Envisat ASAR differential interferometry (Fig.15).
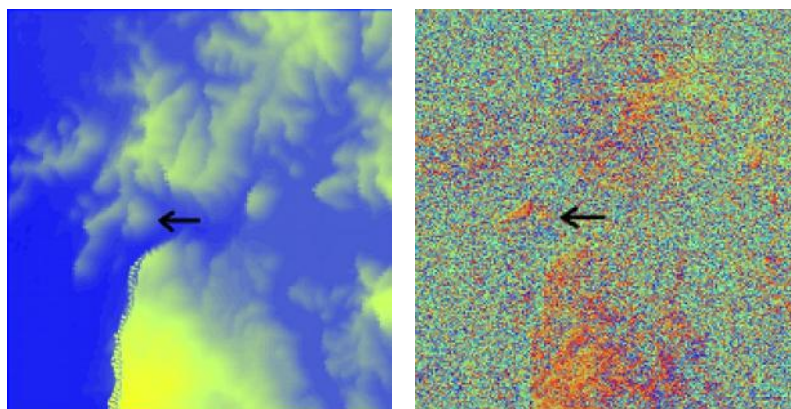
Fig.15: Ragami landslide (black arrow) in Vau Dejes hydroplant from Envisat SAR
images. Left: Digital Elevation Model. Right: DINSAR interferogram.

The landslide looks like a weak fringe in a background of vegetation changes.

3. CONCLUSIONS

Remote sensing satellite images are appropriate for environmental studies. A number of unknown environmental phenomena were identified and some known phenomena spatially and temporary delineated.

The free Internet published LANDSAT and MODIS archives from USGS, as well as Internet archives of European Space Agency ESA accessible in framework of collaborative projects resulted an excellent source for regional studies.

Significant data concerning the spatial and temporal environmental changes were obtained using image processing software for general use combining properly different bands as a single false color image, where different basic colors represent the environmental status in different conditions.

Further complex studies blending together different spatial and temporal environmental parameters from satellite images, calibrated with field data would be necessary.

REFERENCES:

Frashëri N, Çiço B, Paci H, Bushati J. 2010. Using of remote sensing (satellite images) for assessing the environment situation. BALWOIS'2010, Ohrid 25-29 May.

Frashëri N, Beqiraj G, Bushati S. 2013. Environmental Changes of Ohrid-

Prespa from Satellite Imagery. Regional International Conference "The System Prespa Lakes – Ohrid Lake: the Actual State – Problems and Perspectives". Struga-Pogradec, 27-29 October.

Frashëri N, Pano N, Frasheri A, Beqiraj G, Bushati S, Taska E. 2012. A review on anthropogenic impact to the Micro Prespa lake and its damages. European Geosciences Union General Assembly 2012, Vienna Austria, 22-27 April.

Frashëri N, Cico B, Fundo A. 2010. Analysis of Environmental Changes in Shkodra Lake Area Using Remote Sensing. Academy of Sciences of Albania & Academy of Sciences and Arts of Montenegro, International Conference on Shkodra Lake – Status and Perspectives, Podgorica – Shkodra 19-21 June.

Frashëri N, Beqiraj G, Bushati S, Frasheri A, Taushani E. 2014. Remote Sensing Analysis of the Adriatic Shoreline Movements. International Scientific Conference Integrated Coastal Zone Management In The Adriatic Sea, Institute Of Marine Biology, Kotor, Montenegro 29 September – 1 October.

Frashëri N, Frasheri A, Pano N, Bushati S. 2011. Outlook on seawaters dynamics and geological factors of development of Albanian Adriatic coastline. European Geosciences Union General Assembly 2011, Vienna 3-8 April.

Frashëri N, Beqiraj G, Bushati S. 2012. Analysis of aerosols variations using MODIS images. BALWOIS 2012 - Ohrid, Republic of Macedonia, 28 May – 2 June.

Frashëri A, Bushati S, Frasheri N, Dema A. **2014**. Generalized Geophysical Overview of Shkoder-Peje Deep Transversal Fracture. XX Congress of the Carpathian Balkan Geological Association, 24-26 September, Tirana.

Frashëri N, Bushati S, Beqiraj G Lamaj M. 2015. Application of differential inteferometry for analysis of ground movements in Albania. Fringe 2015 Workshop ESRIN, Frascati, Italy, 23–27 March.

Pano N, Frashëri A, Beqiraj G, Frashëri N, Haska H. 2004. Outlook on damages caused from anthropogeneous impact on Micro Prespa Lake. Bulletin of Geological Sciences, Nr. 1, pp.83-93, Geological Survey of Albania.

Pano N, Bushati S, Frashëri N. 2014. Climate Change Impact on Buna. IAEG XII Congress, Torino 15-19 September.

Pano N, Frashëri A, Simeoni U, Frashëri N. 2006. Outlook on seawater dynamics and geological setting factors for the Albanian Adriatic coastline development", Academy of Sciences. *Albanian Journal of Natural and Technical Sciences*. **19-20, (1-2).** Tirana, 152-166.

NOSOCOMIAL BLOOD STREAM INFECTIONS IN PEDIATRIC INTENSIVE CARE UNIT IN ALBANIA

Irena KASMI, Sashenka SALLABANDA
Department of Pediatrics, University Hospital Center, Tirana, Albania
Gentian KASMI
Department of Laboratory University Hospital Center, Tirana, Albania**,**
Eliana IBRAHIMI
Department of Biology, Faculty of Natural Sciences, University of Tirana, Albania

ABSTRACT

The present paper aims to: i) determine the incidence of the most common NI namely blood stream infection (BSI), in the Paediatric Intensive Care Unit (PICU) of University Hospital Center (UHC) and, ii) define the risk factors associated with them and the appropriate prevention and control measures. We performed a prospective study on the incidence of BSI from May 2011 to December 2012 in a single 15 bed-PICU. Six hundred and fifteen patients admitted to PICU, who stayed more than 48 hours were enrolled and monitored for BSI till discharge from ICU or death. Primary BSI was defined as per CDC criteria. Data of patients with BSI was compared with those without BSI. Outcome was measured as length of PICU stay (LOS) and survival or death. In the present study the use of devices (endotracheal tube [ETT], central venous catheter [CVC] and urinary catheter [UC]) was recorded to calculate the device associated infection rate (DAR), incidence density and device utilization ratio (DUR) and risk factors. The results reported: i) 26 episodes of BSI, with an incidence rate 4.3 per 100 admissions or BSI rate 5.7 per 1000 patients-days, ii) 4 clinical sepsis, 7 laboratory confirmed-BSI, 15 episodes of central line associated blood stream infections (CLABSI) among 108 patients using CVC more than 48 hours and, iii) the incidence, incidence density, DAR and DUR among patients using CVC was 14 %, 12 per 1000 patient days, 30.5 per 1000 central line days, 0.39 respectively. The group age mostly affected by CLABSI was children under one year old. The predominant isolates were gram positive (n=16, 73%) of which Staphylococcus aureus was the most common, followed by gram negative (n=6, 27%) of which Enterobacter spp. was the most encountered. Higher frequency and duration of device utilization of CVC, ETT, UC, use of parenteral nutrition, primary surgical diagnosis, younger age, were risk factors for BSI on univariate analysis. The median LOS was longer in patients with BSI compared with those without (17.1 vs.7.05, p=0.0001). However, the presence of BSI, was not associated with a raise in

morta- lity rate which was similar in patients with and without BSI respectively 23.1% vs.17.3% (p = 0.3). BSI rate in our study rate was consistent with that reported by other similar study. The device associated BSI was higher than international standards. This study has established a benchmark for future comparisons. Preventive measures such as improvements in central line insertion and maintenance are necessary to lower this baseline rate in the future comparisons.

Keywords: **nosocomial blood stream infections, children**

1. INTRODUCTION

Intensive care units represent the largest source of infection within the hospital (Donotwitz *et al*, 1982; Donotwitz 1986). The incidence of NI in PICU varies from 6 to 13.7% in developed countries. In developing countries this incidence can be higher, up to 29.6% (Aserbergiene *et al.*, 2009). In particular, device associated infections as blood stream infections (BSI), incidence densities can be up to 19 times higher than in developed countries. Primary Bloodstream Infections (BSIs) are the main Health Care Associated Infections (HAIs) in the USA and the second HAIs in European PICU patients, with substantial effect on cost and outcome (WHO, 2011). Health care associated infections rate are particularly high in neonatal and pediatric ICUs. In studies in neonatal and PICUs in developing countries, CLABSI densities were found to range from 10·2 and 60 (median 18·7) episodes per 1000 catheter-days (Allengrazi *et al*., 2010). Most of the literature concerning BSI is generated from adult studies. Because of significant differences in age, underlying medical conditions and type, and distribution of pathogen these data cannot be directly extrap- olated to children (Cobb *et al.*, 1993).

Although consolidated data on the prevalence of NI in Albanian hospitals is lacking, it is estimated to be in the range of 10-31.6% (Faria *et al*., 2007; Bujari and Kasmi, 2009). Given the importance of the problem, we decided to study the incidence, risk factors, etiological organisms, and outcome of primary nosocomial BSI in our PICU. The results here obtained would be of great importance for the newly created Committee for the Infection Control of the hospital and Prevention already created within the hospital in implementing control measures against this preventable complication of intensive care.

2. MATERIALS AND METHODS

This was a prospective study carried out in the single PICU of the UHC, Tirana, Albania from May 2011 to December 2012. This PICU is a 15-bedded unit having an annual admission average of 450-500 patients with need for invasive and non-invasive treatment. All patients with an ICU stay of >48 h

were involved consecutively into the study. Patients, who died within the first 24hours of admission, stayed less 48 h and those with BSI where the organism was the same as the one cultured prior to PICU admission were excluded from the study. Clinical and laboratory criteria for definition of primary BSI were based on the CDC criteria for infections (Garner, 1996). Criteria applied to patients with >48 h of PICU stay and when the organism cultured was not related to infection at another site. Although all consecutive PICU admissions during the study period that required a CVC, were prospectively studied. In order to esti- mate the true incidence of CLABSIs, the type of CVCs were included: e.g. simple (not tunneled, not antiseptic/antimicrobial) temporary polyurethane catheters. Insertion of temporary CVCs were done under anatomi- cal land mark points and maximum sterile barriers, by the attending physician, using the Seldinger technique. Skin antisepsis was done with povidone iodi- ne 10% and alcohol 70%, due to lack of the recommended 2% chlorexidine gluconate in Albania. No guide wire exchange was performed. Sterile gauze and semi-transparent patches were used as catheter maintenance due to lack of antimicrobial catheter dressing over insertion site. Due to the observational character the study did not require any deviation from routine medical care. The catheters stayed as long as needed if they were functioning without evidence of local or systemic complications.

BSI Definitions

An infection was defined as NI if it occurred 48 hours after admission to the unit or within 48 hours or 48 hours after discharge.

Bacteraemia was defined as the biological documentation of infection, i.e. the result of a positive blood culture. *Laboratory-confirmed bloodstream infection* must meet at least one of the following criteria: i) patient with a recognized pathogen cultured from one or more blood cultures not related to another site, ii) patient with at least *one* of the following signs or symptoms: fever (>38° C), chills, or hypotension *and* at least *one* of the following: common skin contaminant (e.g., diphtheroids, *Bacillus* sp., *Propionbacterium* sp., coagulase-negative staphylococci, or micrococci) is cultured from two or more blood cultures drawn on separate occasions or common skin contaminant (e.g., diphtheroids, *Bacillus* sp., *Propionbacterium* sp., coagulase-negative staphylococci, o micrococci) is cultured from at least one blood culture from a patient with *an intravascular line*, and the physician institutes appropriate antimicrobial therapy.

For patient <1 year of age has at least *one* of the following signs or symptoms: fever (>38° C), hypothermia (<37° C), apnea, or bradycardia *and* at least *one* of the mentioned above. *Clinical sepsis* must meet at least one of the following criteria: Patient has at least *one* of the following clinical signs or symptoms

with no other recognized cause, fever (>38° C), hypotension (systolic pressure <90 mm), or oliguria (<20 cm^3 /hr) *and* blood culture *not* done or *no* organisms or antigen detected in blood *and* no apparent infection at another site *and* physician institutes treatment for sepsis (Garner *et al.*, 2004).

Sampling

Blood from a peripheral vein was drawn for culture, on clinical suspicion of sepsis, starting after first 48 h of ICU stay. The site of sampling was swabbed with povidone iodine. Once swabbed, it was allowed to dry for 2 min. Once dried, it was swabbed with alcohol. Two milliliters of blood was drawn and added 1 ml in each bottle containing 20 ml of media. Samples were delivered to the laboratory within an hour of collection and those collected overnight were stored at room temperature and delivered the next morning.

Data collection

Patients in ward were examined by physicians on duty who are familiar to the phenomenon. *Patient related data,* demographic details, clinical profile (medical, surgical, trauma), place of admission, indication for PICU admission, were recorded in all at the time of enrolment.

Device utilization

Data regarding duration of device utilization (in days) was recorded in all with respect to CVC, mechanical ventilation (MV) and UC. *Invasive procedures and medications data regarding parenteral nutrition, blood transfusion, antibiotic use were recorded.*

Outcome This was measured as length of ICU stay (LOS) and survival or death.

Statistical analysis

The patients for analysis were divided into four age groups: i) infants (aged 1 month to <1 year), ii) preschool children (1-5 years), iii) children (6-12 years) and, iv) teenager (>12 years). NI incidence (number of NI divided by number of patients admitted), NI incidence density (number of NI divided by number of patient-days), device-associated rate (number of NI divided by number of device-days), and device utilization ratio (number of device-days divided by patient-days) were calculated and compared among the age groups. Data are presented as mean ± SD, median and percentages wherever applicable. The Student' s t-test and Mann-Whitney U-test were applied for the parametric data. Chi-square was used for the categorical data. Data of patients with BSI was compared with those without BSI with respect to demographic details,

primary diagnosis surgical, nonsurgical, device utilization and invasive procedures to identify risk factors for BSI. The statistical package used in the study was IBM SPSS version 20.

3. RESULTS

A total 800 admissions were made at PICU from May 2011 to December 2012. 615 children, 321 boys and 294 girls were in the present study involved as they fulfilled the inclusion criteria. The mean age of the study population was 19.4 months (range 0.01-168). There were 22 episodes of primary laboratory confirmed BSI in 22 (3.5 %) patients and 4(0.7%) diagnosed as clinical sepsis; the incidence being 11.2 episodes /1000 patient days. BSI was more common in infants less than one year of age (n=18, 81%) compared with children older than two years of age (n=4, 19%) (Table1).One hundred and eight patients required CVC insertion. Among them, 15 episodes of CLABSI were confirmed. The mean rate of CLABSI was 30.5 per 1000 central-catheter-days with a catheter utilization rate of 39% (Table 2). Of the 15 CLABSIs, 12 isolates were gram positive bacteria, S.aureus (n=10,45%), CoNS (n=2, 9%),3 isolates were gram negative, Enterobacter spp.(n=1,4.5%), Klebsiella pneumonia (n=1,4.5%), no identified gram negative (n=1,4.5%). Among the 7 laboratory confirmed BSI episodes 4 isolates were gram positive, CoNS (n=3,13.5%), Staphylococcus MRSA(n=1,4.5%), 3 isolates were gram negative, Enterobacter sp.(n=2, 9%), Escherichia coli (n=1, 4.5%).

Table 1 Characteristics of 615 patients in the PICU with and without BSI during the study period

Characteristics	Total (N =615)	With BSI (N=26)	NO BSI=589	P value
Age (months) Mean (min max)	19.4(0.01;168)	19.9(0.01;144)	19.3(0.01;168)	0.92
Sex n (%)				
Male	321(52.2)	13(50)	308 (52.3)	
female	294(47.8)	13(50)	281(47.7)	
1mo- <1 year	377(61.3)	20(80)	357(60.7)	
2-5 years	185(30.1)	4(15.3)	181(30.7)	
5-12 years	35(5.7)	0(0)	35(5.9)	
>12 years	18(2.9)	2(7.7)	16(2.7)	
Primary diagnosis at admission, n (%)				
Medical	437(71)	10(38.5)	427(72.5)	
Surgical	178(29)	16(61.5)	162(27.5)	0.001
Emergency department	450(73.1)	10(38.5)	440(74.7)	
Others wards	165(26.9	16(61.5)	149(23.7)	

Irena KASMI, Sashenka SALLABANDA,
Gentian KASMI, Eliana IBRAHIMI

Device utilization, CVC n (%)	108(17.5)	15(68.2)	93(15.7)	0.001
	108(17.5)	15(68.2)	93(15.7)	
Type of catheter,CVC n	0(0)	0(0)	0(0)	
Non tunnelled simple				
Permanent				
	111(18)	20(76.9)	91(81.9)	0.001
Parenteral nutrition, n(%)	208(33.8)	24(92)	184(88.5)	0.001
Blood transfusion, n(%)	345(56)	19(73)	326(55.3)	0.005
ETT, n(%)	339(55)	22(84.6)	317(53.8)	0.001
CU, n(%)				
	108(17.5)	6(23)	102(17.3)	0.43
Outcome Death n, (%)	0.8(± 2.3)	5.5(±4.9)	0.59(±1.96)	0.001
Line duration mean (SD)	7.47(± 6.1)	17.1 (±8.6)	7.5(±5.68)	0.001
PICU stay(days)mean (SD)				

Table 2 CLABSI incidence rates and device utilization ratio in the age groups

Age groups	Bloodstream nosocomial infections				Central line
	Number of NIs (n)	Incidence			Device utilization Ratio
	Incidence	Device	Density	associated rate	
1 month - <1 year	13.0	18.5	15	40.6	0.37
1-5 years	2.0	5.3	6	17.5	0.37
6-12 years	0.0	0.0	0.0	0.0	0.38
>12 years	0.0	0.0	0.0	0.0	0.81
All	15.0	14.0	12	30.5	0.39

Further analyses of the pathogens showed 16(73%) to be Gram-positive, with *Staphylococcus aureus* as the leading causative pathogen (N=11, 50%) followed by Coagulase-negative staphylococci (n=5, 22.8%). Among Gram-negative Enterococcus spp. were the leading causative pathogens, responsible for 3 (13.5%) of the 6 (27%) cases in the Gram-negative group (Table 3).

Table 3 Etiological organisms for BSI (n=22)

Staphylococcus aureus (S.aureus)	10(45.5)
Staphylococcus aureus (MRSA)	1(4.5)
Staphylococcus epidermidis(CONS)	5(22.8)
Enterobacter spp.	3(13.7)
Escherichia coli (E.coli)	1(4.5)
Klebsiellapneumoniae (Kleb)	1(4.5)
Others gram negative	1(4.5)

Risk factors associated with BSI included younger age, primary diagnosis (surgical, medical; *P<0.001*), parenteral nutrition *(P<0.001),* the use of ETT

(P=0.005), the use of CU *(P<0.001)*. Variables significant on univariate testing (*P* < 0 .05) were entered in multiple logistic regression analysis. Patients who had parenteral nutrition infused through the line and blood product transfused were more likely to develop BSI (adjusted OR respectively: [OR= 7.9; 95% CI: 2.6-19.6; p=0.031]; [OR= 10.9; 95% CI: 2.04-55.5; p=0.005]). Univariate analysis of the outcome of BSI revealed that patients with BSI had median LOS longer compared with those without (17.1 vs. 7.5 days, p=0.0001). The crude mortality rate associated with BSI was 23.1% compared with 17.3% in patients without BSI (p = 0.3).

4. DISCUSSIONS

Carried out for the first time in the PICU of UHC, Tirana, Albania, this prospective study aimed at evaluating the rate, risk factors and outcomes of BSI among pediatric patients using the NI CDC definitions (Garner *et al*., 1996). In addition, it is an appropriate means to address for future studies in support of the newly created committee of infection control and prevention within the hospital and allow for comparison of international data. The mean rate of CLABSI in our PICU (30.5 per 1000 central-catheter-days) was very high compared with the mean rate of CLABSI reported by systematic review conducted by WHO which varies from 2.1-4.3 in developed countries to 10.2-60 episodes per 1000 catheter days in developing countries (Pessoa-Silva *et al*., 2004; Allengrazi *et al*., 2010; WHO 2011). Richard et *al.,* (1999) examined BSI in 61 PICUs in the USA and found it to be the most common nosocomial infection with a rate ranging from 0.0 to 20/1000 catheter-days. Stover et *al.,* (2001) reported PICU BSI rates from 23 US hospitals with wide variations in the rates ranging from 0.0 to 18.5/1000 catheter-days. This variability depends on unit related parameters such as size and settings, patient related parameters such as severity and type of illness, site and conditions under which the catheter was placed. This data are missing in this study. It is difficult to explain the reason for our high rate as there are no previous data with which to compare. The source for the high rate is partly the features of our PICU's population and the variability in critical care practices and local infection control. Firstly, the majority of our patients were young infants (61%). It is well known that CLABSI is more common in young infants (Chodekar *et al*., 2011). It is well demonstrated in more than one study the high rate of CLABSI in NICU patients (NNIS, 2001; 2003). In the USA, PICUs have more than 50% of postoperative cardiothoracic patients who are quite stable, with less illness severity and shorter length of stay, opposite to our population with postoperative patients being 29%, high medical emergency patient being71%. O'Grady *et al*., (2002) reported that trauma patients documented the highest CLABSI rate, and non-elective

admission for medical management was found as independent predictor of CLABSIs in the study of Costello *et al.,* (2009) supporting our explanation. Secondly, the lack in nursing staff being of a ratio nurse: patient 1.5-2:1during the day and 3-3.5: 1 during night shift may be another reason as is well known that understaffing in ICU is an independent risk factor for NI in general and specifically for BSI which is highly dependent on hygienic behavior (Robert *et al*., 2000; Schwab *et al*., 2012). Thirdly, the lack of consistency in application of written guidelines for the access and care of CVCs in our PICU, the use of antimicrobial-impregnated catheters, the implementation of annual hospital wide hand washing campaigns, the creation of physical barriers between patient beds, the use of 2% chlorexidin skin disinfectant for all central lines for all patients of all ages have, may have contributed to the high rate. Finally, although it was intended to use the CDC NNIS definition for CLABSI, this was not strictly applied. In our study, we did not require evidence that the catheter was the focus of infection. It is therefore possible that some of the episodes attributed to the presence of CVC, could have originated from the gastrointestinal tract or elsewhere.

The distribution of causative micro-organisms for BSI in our PICU is similar to that reported in the NNIS data (NNIS, 2001; 2003). Gram-positive organisms, were predominant (73%) in our population, with S.aureus (27%) being the leading pathogen, whereas Gram-negative organisms were less encountered (27%) of which Enterobacter spp. was the most encountered.

In multiple logistic regression analysis parenteral nutrition infusion and blood product transfusion through the line were found to be a risk factor for CLABSI in our PICU as supported by similar study (Almuneef *et. al*., 2006; Wylie *et. al*., 2008). This is supported by Singh-Naz and colleagues (1996) who noted that children receiving parenteral nutrition and antimicrobial therapy were associated with the higher risk of developing nosocomial infection. It has been emphasized in our PICU that infusion of fluid emulsion should be completed within 12 h, and that of blood or other blood products within 3 h. Prior antibiotic therapy was not studied as a risk factor for BSI in our study, during which approximately 98% of the enrolled patients and 100% of them with CVC were taking antibiotics prior to its development, although it has been mentioned in studies that receiving antibiotic for more than 10 days could be a risk for developing nosocomial infections including lower respiratory tract, BSI, and urinary tract (Andrea, 2005). The severity of illness of our patients was not reported since the PRISM III score, which is now in use, was partially in use during the study period. A number of knowing risk factor for BSI as the place of insertion, the anatomical site, the catheter insertion necessity, factors that should be taken in consideration for future studies.

The mean line duration and the PICU length of stay were longer in the BSI

cases. The duration of catheterization is reported as independent infection risk factor in some studies, a fact that was confirmed by our results, showing longer duration of catheterization in patients with catheters who developed CLABSI (Wylie *et al*., 2008; Rey *et al*., 2011). However, the presence of BSI, under the conditions of our study, was not associated with the mortality rate which was identical in patients with and without BSI (23.1% vs.17.3% (p = 0.3), in agreement to the majority of published data (Smith 2008; O'Grady *et al*., 2011).

The major limitation of our study relates to a relatively small study center, which poses barriers in the generalization of our results. Although, there are strong points as well because international definitions we used enabled us to compare our data to international standards and allow us to set a benchmarking for future comparisons.

5. CONCLUSIONS

BSI especially CLABSI rate in our study is higher than the international standards. We set the baseline for CLABSI rate, causative microorganisms and risk factors as indicative infection control measures to be taken.

REFERENCES:

Allegranzi B, *S*epideh B, Combescure Ch, Graafmans W, Attar H, Donaldson L, Pittet D. 2010. Burden of endemic health-care-associated infection in developing countries: systematic review and meta-analysis. WHO.

Almuneef MA, Memish ZA, Balkhy HH, Hijazi O, Cunningham G, Francis C. 2006. *Rate*, *risk factors* and *outcomes* of *catheter-related bloodstream infection* in a *pediatric intensive care unit* in *Saudi Arabia. Journal of Hospital Infection*, **62(2)**:**207–213.**

Andrea M. 2005. Pediatric catheter-related bloodstream infections latest strategies to decrease risk. *AACN Clinical Issues*, **16(2):** 185–198.

Asembergiene J, Gurskis V, Kevalas R, Valinteliene R. **2009**. *Nosocomial infections in the pediatric intensive care units in Lithuania. Medicina (Kaunas)*, **45(1);** 29-36.

Bujari I, Kasmi G 2009. Incidence of nosocomial infections in the pediatric Intensive Care Unit, University Hospital Center, Tirana. *Albanian Journal of Natural & Technical Sciences* (*AJNTS*). July, **26** (**2**), 121-126.

Cobb JP, Danner RL, Saunders WB.1993. Nosocomial infections in the practice of pediatric critical care. Textbook of Pediatric Critical Care, Philadelphia: Holbrook. 868-82.

Chopdekar K, Chande C, Chavan S, Veer P, Wabale V, Vishwakarma

K, Joshi A. 2011. Central venous catheter-related blood stream infection rate in critical care units in a tertiary care, teaching hospital in Mumbai. *Indian Journal of Medical Microbiology*; **29(2)**: 169-171.

Costello JM, Graham DA, Morrow DF, Potter-Bynoe G, Sandora TJ, Laussen PC**. 2009**. Risk factors for central line- associated bloodstream infection in a pediatric cardiac intensive care unit. *Pediatric Critical Care Medicine*, **10**:453-459.

Donowitz LG, Wenzel RP, Hoyt JW. 1982. High risk of *hospital-acquired infection* in the *ICU patient. Critical Care Medicine, 10:355–357.*

Donowitz LG. 1986. High risk of *nosocomial infection* in the *pediatric critical care patient. Critical Care Medicine, 14:26-28.*

Faria S, Sodano L, Gjata A, Dauri M, Sabato AF, Bilaj A, Mertiraj O, Llazo E, Kodra Y, Schinaia N; Prevalence Study Group. 2007. The first prevalence survey of nosocomial infections in the University Hospital Centre 'Mother Teresa' of Tirana, Albania. *Journal of Hospital Infection*, **65**: 244–250.

Garner JS, Jarvis WR, Emori TG, Horan TC, Hughes JM. 1996. CDC definitions for nosocomial infections. Olmsted RN, ed. APIC Infection Control and Applied Epidemiology: Principles and Practice. St. Louis: Mosby; A-1—A-20.

National Nosocomial Infections Surveillance (NNIS). 2001. System report, data summary from January 1992–June, issued August 2001. Elsevier. *American Journal of Infection Control*, **29**: 404-421.

National Nosocomial Infections Surveillance (NNIS). 2003. System report, data summary from January 1992 through June, issued August 2003. Elsevier. *American Journal of Infection Control*, **31**: 481-498.

O'Grady NP, Alexander M, Dellinger EP, **Gerberding JL, Heard SO, Maki DG, Masur H, McCormick RD, Mermel LA, Pearson ML, Raad II, Randolph A, Weinstein RA**. 2002. Guidelines for the prevention of intravascular catheter-related infections. *Pediatrics*, **110**:1-24.

O'Grady NP, Alexander M, Burns LA, O'Grady NP, Alexander M, Dellinger EP, **Garland J, Heard SO, Lipsett PA, Masur H, Mermel LA, Pearson ML, Raad II, Randolph A, Rupp ME, Saint S, and the Healthcare Infection Control Practices Advisory Committee (HICPAC). 2011**. Healthcare Infection Control Practices Advisory Committee (HICPAC). Guidelines for the prevention of intravascular catheter- related infections. *Clinical Infection Disease*, **52**:162-93.

Pessoa-Silva CL, Richtmann R, Calil R, Santos RMR, Costa MLM, Frota AC, Wey SB. 2004. Healthcare-associated infections among neonates in Brazil. *Infection Control and Hospital Epidemiology*; **25(9)**: 772-7.

World Health Organization (WHO). 2011. *Report on the Burden of Endemic Health Care-Associated Infection Worldwide.*

Rey C, Alvarez F, De-La-Rua V, Concha A, Medina A, Díaz JJ, Menéndez S, Los-Arcos M, Mayordomo-Colunga J. 2011. Intervention to reduce catheter-related bloodstream infections in a pediatric intensive care unit. *Intensive Care Medicine*, **37**:678-85.

Richard MJ, Edwards JR, Culver DH, Gaynes RP, the National Nosocomial Infections Surveillance System. 1999. Nosocomial infections in pediatric intensive care units in the United States. *Pediatrics*, **103 (4)**:e39.

Robert J, Fridkin SK, Blumberg HM, Anderson B, White N, Ray SM, Chan J, Jarvis WR. 2000. The influence of the composition of the nursing staff on primary bloodstream infection rates in a surgical intensive care unit. *Infection Control & Hospital Epidemiology,* **21(1),** 12-17.

Schwab F, **Meyer E**, **Geffers C**, **Gastmeier P**. **2012**. Understaffing, overcrowding, inappropriate nurse: ventilated patient ratio and nosocomial infections: which parameter is the best reflection of deficits? *Journal of Hospital Infection*. **8(2)**:133-9.

Singh-Naz N, **Sprague BM**, **Patel KM**, **Pollack MM. 1996.** Risk factors for nosocomial infection in critically ill children: a prospective cohort study. *Critical Care Medicine*, **24**: 875-878.

Smith JM. 2008. Catheter-related bloodstream infections in children. *American Journal of Infection Control*, **36**: S173- S173.

Stover BH, Shulman ST, Bratcher DF, Brady MT, Levine GL, Jarvis WR, Pediatric Prevention Network. 2001. Nosocomial infection rates in US children's hospitals' neonatal and pediatric intensive care units. *American Journal of Infection Control*, **29**; 152—157.

Wylie MC, Graham DA, Potter-Bynoe G, Kleinman ME, Randolph AG, Costello JM, Sandora TJ. 2008. Risk factors for central line-associated bloodstream infection in pediatric intensive care units. *American Journal of Infection Control*, **36(10):** S173.

REIRRADIATION OF PATIENTS WITH BONE METASTASIS

Elvisa KOZMA and Agim SALLAKU

University Hospital Centre, Mother Theresa, Tirana, Albania

ABSTRACT

The present paper aims to evaluate the clinical outcome of reirradiated patients with bone metastasis in terms of pain improvement, local control, and development of myelopathy. 20 patients with bone metastasis who complained pain have been reirradiated from 2012-2014. The median interval between the two irradiations was 17 months (range 10-38 months). Karnofsky performance score at the time of retreatment ranged from 50% to 80%, and median age was 46 years (range 26-74). The first radiation dose varied from 30 Gy to 50 Gy with daily fraction doses of 1.8-3Gy. The total dose of retreatment ranged from 8-30 Gy with daily fraction doses 3-8 Gy, and the cumulative dose ranged from 38 Gy to 70 Gy (median 53 Gy). All patients were followed up prospectively and all of them tolerated well reirradiation. The median follow up after the last treatment was 12 months (range 8-26 months). Here, 12 patients out of 20 had complete pain relief and the remainder had partial pain relief. At the last follow up only one patient was still alive and the others were dead from systemic disease progression. We didn't detect any serious acute side effect and no patient showed treatment-induced neurologic abnormalities. The results reported that reirradiation of spinal cord within the dose range we used is a feasible treatment that provides clinical benefits and improvements of quality of life.

Key words: reirradiation, spinal cord, myelopathy

1. INTRODUCTION

Nowadays improvements of cancer therapy have led to better survival rates and an increased incidence of recurrences in site of previous conventional radiotherapy. Many of them need to be reirradated as the only treatment option and are not candidates for a surgical approach because of the limited indications (a poor performance status, co-morbidity, presence of other sites of metastases or localization of the metastasis in the skeleton).

However, reirradiation still remains a challenge for the radiation oncologist as a result of limited evidence. The right interval between two irradiations, fractionation and total dose remain unclear.

As spinal cord is a dose-limiting organ and radiation can induce damage at even a small volume of it and impair normal function, we should be very careful not to lead to myelopathy, which is irreversible and can radically reduce the quality of life of the treated patient. Radiotherapy is a well-accepted treatment modality for localized pain with a 60–80% likelihood of overall pain relief (Wu *et al.,* 2003), however higher RT doses, larger doses per fraction, and previous exposure to radiation could be associated with a higher probability of developing radiation induced myelopathy (Makbeth *et al.,* 1996). Experimental data report that the total dose of the first and second radiotherapy, interval to reirradiation, length of irradiated spinal cord, and age of the treated animals influence the risk of radiation induced myelopathy (Rades *et al.*, 2008).

The present paper investigates the feasibility of spinal cord reirradiation, its effect on clinical outcome and improvement of quality of life.

2. MATERIAL AND METHODS

In the present investigation 20 patients with bone metastasis have been reirradiated from January 2012 until December 2014. 12 patients out of 20 were men and the reminder were women with a median age of 54 years (range 26-74 years). Except one all other patients have only bone metastasis at the time of retreatment. They have received multimodality treatment before and also have used corticotherapy. Three patients were reirradiated to the cervical spine, twelve to the thoracic spine and five to the lumbar spine. Patients were treated for different primary tumors (see tab.1), but the reason of the second irradiation was pain progression and refractory to analgesics.

The total dose of the initial radiation therapy ranged from 30 to 60 Gy, with single doses 1.8–3Gy. All patients had initial multifractionated radiotherapy treatment and sixteen of them were treated with a Cobalt 60 machine and four with a Linear accelerator using a three-dimensional conformal irradiation technique. The median interval between two irradiations was 17 months (range 10-38 months) and the total dose of reirradiation ranged from 8 to 30 Gy, with daily fraction of 3–8Gy. Ten patients received single dose radiotherapy and 10 others multifractionated treatment (6 patients 5x4Gy and four patients 10x3Gy). Twelve patients were retreated by a Cobalt 60 machine and eight patients by a Linear accelerator using a three-dimensional conformal irradiation technique.

Table 1. Patients' characteristics

Characteristic	Number of patients
Age	
50years	7
50years	13
Gender	
Female	12
Male	8
Type of tumor	
Breast cancer	6
Prostate cancer	4
Lung cancer	3
Rectal cancer	3
Renal cancer	2
Rectal cancer	2
Lymphoma	1
Visceral metastasis at the time of retreatment	
Yes	
No	1
	19
Vertebral position	
Cervical	3
Thoracal	12
Lumbar	5
Time between two RT treatments	
12months	6
12months	14
Radiation schedule	
Single dose	10
Multiple fractions	10

3. RESULTS

Reirradiation was tolerated well by all patients. No acute side effects $\geq$ grade 2 was seen. One patient developed dysphagia G1 and another diarrhea G1.

After the second month of retreatment twelve patients have complete pain relief and eight others have only partial pain relief. The evaluation of pain was done based on patient self assessments of pain using a numerical rating score from 0 (no pain) to 10 (worst imaginable pain). Also every change in pain medication during follow-up was recorded.

Afterwards we followed up the patients every three months. In the next 6 months, no one of twelve patients complain pain as a major symptom. The

median follow up after the last treatment was 12 months (range 8-26 months). At the last follow up all patients except one who is still alive complained pain, but the cause of death in all patients originated from progressive disease.

No patient showed any neurologic abnormalities affecting motor or sensory function. They were all able to walk and were continent for urine and stool until the time of death.

4. DISCUSSION

Reirradiation of spinal cord is still under investigation. Radiation oncologists are very reserved about it as a result of existing concerns that reirradiation may induce myelopathy and severe neurologic dysfunction. Symptoms may develop months until years after reirradiation (Ryu *et al*., 2000). Nevertheless actually exists prospective and retrospective data that supports retreatment of painful bone metastasis following initial irradiation (Grosu *et al*., 2002, Maranzano *et al*., 2011).

There are three clinical situations when reirradiation can be used: in case of no response, after partial response from first irradiation and hope to achieve further pain relief, and in case of recurrence of pain during follow up after complete or partial response with initial radiotherapy.

In our case all patients were reirradiated for progressive pain after partial or complete response with initial radiotherapy.

Although, literature shows a 12-months local-recurrence rate of 39% with single dose radiotherapy and of 23% for patients treated with fractionated courses (Rades *et al*., 2009) many radiation oncologist use single doses for reirradiation to avoid side effects. In the present investigation single doses were used in 10 patients. The dose of second irradiation was also conditioned by patient's conditions and their life expectancy. Patients with a poor expected survival can benefit from a shorter course of radiation because a shorter overall treatment time means less daily trips to the radiation oncology department and positioning on the treatment couch and as a result less efforts for the patient and less costs for the treatment (Van den Hout *et al*., 2003).

During reirradiation are used different daily fractionations and different total doses which make difficult the comparison between their effects in spinal cord. We should convert the given physical dose to biological equivalent dose (BED), in order to make possible the comparison between them.

Reirradiation has a higher risk of causing myelopathy because the cumulative BED to spinal cord is higher (Maranzano *et al*., 2001, Wong *et al*., 1994).

Nieder and colleagues (2006) reported that clinical data on spinal cord reirradiation allowed the development of a risk score that distinguish patients who may develop radiation induced myelopathy on the basis of three variables

which are cumulative BED, BED of each single radiotherapy course, and interval between the courses. The risk of radiation induced myelopathy appeared small after cumulative.

BED $\leq$ 135.5 Gy$_2$, when the interval from the first RT and reirradiation is not shorter than 6 months, and the BED of each course was $\leq$ 98 Gy$_2$ (Nieder *et al*., 2006).

In our investigation we prescribed radiotherapy regimens which cummulative BED did not exceed 150 Gy$_2$, and BED of each course was $\leq$ 98 Gy$_2$ to the tumor to the tumor and $\leq$ 75 Gy$_2$ to the spinal cord. All our patients have an interval between the courses more than 6 months. So, according to Nieder's risk categories the majority of our reirradiated patients had an estimated low risk of radiation induced myelopathy and indeed no one of them developed myelopathy.

Our data supports the retreatment of patients with pain from bone metastasis following initial irradiation, especially after initial response. The three types of fractionation we used (single dose 8Gy, 5x4Gy and 10x3Gy) were safe and effective for pain relief, but patients who were treated with multifractionated radiotherapy showed more complete pain relief than patients treated with single dose reirradiation (ten from twelve patients). Nevertheless, we should take into account that we treated a small number of patients from which clear conclusions cannot be drawn.

5. CONCLUSION

Reirradiation of spinal cord using the fractionations and total doses we used is a safe treatment with very good palliative effects and no risk of complications for the rest of life.

As the number of patients here in the present investigation involved was low, evaluating the fraction size or the total dose with optimal clinical effects and minimal side effects was impossible.

REFERENCES

Grosu AL, Andratschke N, Nieder C, Molls M. 2002. Retreatment of the spinal cord with palliative radiotherapy. Official journal of the American Society for Radiation. *International Journal of Radiation Oncology Biology Physics,* **52**:1288–92.

Macbeth F, Wheldom TE, Girling DJ, Stephens RJ, Machin D, Bleehen NM, Lamont A, Radstone DJ, Reed NS. 1996. Radiation

myelopathy: estimates of risk in 1048 patients in three randomised trials of palliative radiotherapy for non-small cell lung cancer. The Medical Research Council Lung Cancer Working Party. *Clinical Oncology*, **8**:176–181.

Maranzano E, Bellavita R, Floridi P. 2001. Radiation-induced myelopathy in long-term metastatic spinal cord compression patients after hypofractionated radiotherapy: a clinical and magnetic resonance imaging analysis. *Radiotherapy and Oncology*, **60**:281–288.

Maranzano E, Trippa F, Casale M, Anselmo P, Rossi R. 2011. Reirradiation of metastatic spinal cord compression: Definitive results of two randomized trials. *Radiotherapy and Oncology*, **98**; 234–237.

Nieder C, Grosu A L, Andratschke NH, Molls M. 2006. Update of human spinal cord reirradiation tolerance based on additional data from 38 patients. Official journal of the American Society for Radiation. *International Journal of Radiation Oncology Biology Physics* **66, (5);** 1446–1449.

Rades D, Lange M, Veninga T, Rudat V, Bajrovic A, Stalpers LJ, Dunst J, Schild SE. 2009. Preliminary results of spinal cord compression recurrence evaluation (score-1) study comparing short-course versus long-course radiotherapy for local control of malignant epidural spinal cord compression. Official journal of the American Society for Radiation. *International Journal of Radiation Oncology Biology Physics,* **73**:228–34.

Rades D, Rudat V, Veniga T, Hoskin P. 2008. Prognostic factors for functional outcome and survival after reirradiation for in-filed recurrences of metastatic spinal cord compression. *Cancer*; **113**:1090–1096.

Ryu S, Gorty S, Kazee AM, Bogart J, Hahn SS, Dalal PS, Chung CT, Sagerman RH. 2000. Full dose reirradiation of human cervical spinal cord. *American Journal of Clinical Oncology*, **23**:29–31.

Van den Hout WB, Van der Linden YM, Steenland E, Wiggenraad RG, Kievit J, de Haes H, Leer JW, 2003 Single- versus multiple-fraction radiotherapy in patients with painful bone metastases: Cost-utility analysis based on a randomized trial. *Journal of National Cancer Institute*, **95**:222-229.

Wu JS, Wong R, Johnston M, Bezjak A, Whelan T, 2003 Meta-analysis of dose-fractionation radiotherapy trials for the palliation of painful bone metastases. Official journal of the American Society for Radiation. *International Journal of Radiation Oncology Biology Physics,* **55**:594–605

Wong CS, Van Dyk J, Milosevic M, Laperriere NJ, 1994 Radiation myelopathy following single courses of radiotherapy and retreatment. Official journal of the American Society for Radiation. *International Journal of Radiation Oncology Biology Physics,* **30**:575–81.

AN EFFECTIVE HYBRID OF BAT ALGORITHM AND HILL CLIMBING FOR GLOBAL OPTIMIZATION OF HIGH-DIMENSIONAL FUNCTIONS

Ali Osman TOPAL

Information and Communication Technologies Coordinating Office, Epoka University, Tirana, Albania

Oguz ALTUN

Yildiz Technical University, Istanbul, Turkey

Yunus Emre YILDIZ

Information and Communication Technologies Coordinating Office, Epoka University, Tirana, Albania

ABSTRACT

The present paper provides information about a hybrid optimization algorithm which is a combination of Bat Algorithm (BA) and Hill Climbing (HC). This algorithm incorporates the diversification ability of BA with intensification ability of HC. The hybrid optimization algorithm is verified by testing on a large set of numerical test functions and results obtained were compared with the results obtained by bat algorithm (BA) and three state-of-the-art BA variants. The results reported that the performance of the proposed method is better than the bat algorithm. In addition, it is comparable to three state-of-the-art modified BA algorithms in terms of the quality of final solution and its convergence rates for high dimensional problems.

Keywords: Bio-inspired algorithm, Numerical optimization, Bat algorithm, Hybrid bat algorithm.

1. INTRODUCTION

Many nature-inspired algorithms have been invented over the past a few decades and applied with success to many numerical and combinatorial optimization problems. The aim of the optimization algorithms is to find a set of independent variables which are maximizes or minimizes the value of dependent variables. Ant Colony Optimization (Dorigo *et al*., 1996), Particle Swarm Optimization (Kennedy and Eberhart 1995), Artificial Bees Algorithm (Karaboga 2010), Bat Algorithm (Yang 2010), and Artificial Immune System

(Kim *et al*., 2007) with increasing popularity are examples of nature inspired algorithms. Since the real-world optimization problems are getting more complicated, higher dimensioned and more dynamic, it follows that meta-heuristics algorithms' popularity will continue to increase, as the finding the optimal solution becomes very difficult within a reasonable amount of time.

Bat Algorithm (BA) is one of the nature-inspired algorithms recently presented and successfully applied to solve numerous optimization problems in diverse fields. It is a population-based search algorithm that was based on the echolocation behavior of bats (Yang 2010). BA, like Yang's previous algorithms; Cuckoo Search (Yang and Deb 2009) and Firefly (Yang 2009), combines the advantages of existing algorithms, especially Particle Swarm Optimization and Harmony search. Many researches on BA have been carried out recently, due to its simplicity, convergence speed and population feature such as: i) multi-objective bat algorithm (MOBA) (Yang 2011), ii) self -adaptive modified bat algorithm (SAMBA) (Baziar *et al*., 2013), iii) local memory search bat algorithm (LMSBA) (Yuanbin *et al*., 2013) and, iv) adaptive bat algorithm (ABA) (Wang *et al*., 2013). BA shows good results when dealing with lower-dimensional optimization problems, but its performance decreases for high-dimensional problems as it becomes more exploitative very fast and may stack in a local optimum.

Hill Climbing (HC) is related to gradient ascent. Although good for unimodal problems, HC doesn't guarantee the best possible solution for multi-modal problems as it is easily trapped in a local optimum when many local optima could be met. Many researchers (Renders and Bersini 1994; Lim *et al*., 2006; Yin *et al*., 2006; Yildiz 2009) prefer hybrid algorithms with hill climbing to increase the intensification of the search due to its simplicity and its high reliability for finding local optimum.

This paper proposes a hybrid bat algorithm with hill climbing to increase the local search of BA. The hybrid is named as BAHC. The performance of the BAHC is extensively evaluated on eleven benchmark functions from Yao *et al*., (1999) and compared with standard Bat Algorithm and three state-of-the-art modified BA algorithms.

The remainder of the paper provides information about the Bat Algorithm, its variants used to compare the performance of BAHC, HC, the hybrid BAHC and experimental results and comparison with other algorithms on optimization test functions. At the end some conclusions are drawn.

2. BAT ALGORITHM (BA)

Bats have unique ability to detect insects and avoid obstacles by using a high frequency sound based system which is called echolocation. Bats emit sound waves and listen to the returning echoes and they can draw 3D picture of their

surrounded environment in their brain even in pitch black dark. Bats can tell the shape, size, texture of a tiny prey, in which direction the prey is heading and even the speed of the prey by using the delay time and loudness of the response and the time difference between the ears. Bats have also the ability to change the way they emit the sound pulses (Richardson 2008; Gandomi *et al.*, 2013; Topal and Altun 2014).

The algorithm starts with the n bats being placed randomly in the search space. Velocity v_i, frequency f_i, pulse rates r_i, and loudness A_i are initialized for each bat at the beginning. Bat's position and velocity are updated by using (1, 2, and 3).

$$f_i = f_{min} + (f_{max} - f_{min}) \; \mathsf{S}, \tag{1}$$

$$V_i^t = V_i^{t-1} + (x_i^t - x^*)f_i, \; , \tag{2}$$

$$x_i^t = x_i^{t-1} + V_i^t \tag{3}$$

where $\mathsf{S} \in [0,1]$. Then the algorithm evaluates the finesses (solutions) and chooses the current global best position x^* at time step t. After these updates, if the bat's pulse rate is low (it means bat is far away from the prey), with a high probability it will fly near the current best bat and make a random short fly there. If its pulse rate is high then it should be near prey and it will again with a high probability just make a random fly around its current position

$$x_{new} = x_{old} + \in A^t, \tag{4}$$

where $\in \in [-1, 1]$ is a random number, while $A^t = < A_i^t >$ is At $= 1/N \sum_{i=1}^{N} A_i^t$ the average loudness of all bats at this time step, where N is the number of bats. After this fly if the bat's position is better than the current global best and its loudness is loud enough to be greater than a random number, the bat will fly to this position and current global best will be updated with the new one. The bat's pulse rate r_i will be increased and loudness A_i will be decreased. Then again finesses will be evaluated and the current global best x^* will be found.

3. MODIFIED VERSIONS OF BAT ALGORITHMS

Here, the theory of three state-of-the-art modified versions of BA algorithms: Chaotic Local Search-based Bat Algorithm (CLSBA), Adaptive Bat Algorithm (ABA), and Novel Adaptive Bat Algorithm (NABA) is briefly reported.

3.1 Chaotic Local Search-based Bat Algorithm (CLSBA)
CLSBA (Chandrasekar, 2013) is a combination of the standard BA with

chaotic sequences generated by the logistic map. The use of chaos makes the frequency adaptive and more random in nature to balance the trade-off between exploration and exploitation. The frequency of pulse emission is modified as follows in CLSBA:

$$f_i(t + 1) = \sim * f_i(t) * (1 - f_i(t)) \qquad (5)$$

While the step size of random walk parameters $\in$ (Eq. 4) varies within the range [-1, 1] in BA, it varies within the range [-SF(t), SF(t)] in CLSBA. SF is the scaling factor which changes dynamically according to the Rechenberg's mutation rule (Rechenberg 1994). If the value of SF(t) is low, exploitation increases. If value of SF(t) is high, exploration accelerates.

3.2 Adaptive Bat Algorithm (ABA)

Wang and colleagues (2013) introduced an improved bat algorithm to solve BA's premature convergence problem. In ABA, Wang and colleagues improved the random walk equation, but they also modified the frequency and the velocity equations in BA where each bat uses the same frequency increment for the velocity which makes bat's flight behavior lack of flexibility. In ABA, Wang proposed a new method to let each bat dynamic and adaptively adjust its flight speed and its flight direction. ABA is targeting to increase the speed of the bat which is farther from the prey. The farther the distance between the bat and its prey (global best solution), the faster the speed flying to its prey.

Secondly, Wang improved the random walk by combining it with shrinking search. Bats which are far away from the current best, make random fly near to the current best. The step size of random fly within the range [-1, 1] in BA while it shrinks in ABA as the iteration proceeds. ABA targets to increase the intensification of the search by shrinking search method.

3.3 Novel Adaptive Bat Algorithm (NABA)

NABA (Kabir *et al.*, 2014) incorporates two techniques within BA, which include the Rechenberg's 1/5 mutation rule (Rechenberg, 1994) and the Gaussian/Normal probability distribution to produce mutation step size.

NABA offers new equation to generate new solution rather than doing just random walk. It modifies the random walk equation (Eq.4) in line 12 of original Bat algorithm. NABA controls the random walk step size by the variance of Gaussian/Normal distribution. The modified equation is as follows:

$$x_{new} = x_{old} + \in A^t N(0, \ ) \qquad (6)$$

where $\dagger$ is the standard deviation.

Kabir used the Rechenberg's 1/5 mutation rule (Rechenberg 1994) to adaptively change the random walk step size and pulse rate to control the exploration and exploitation. If the success-rate is less than 1/5, NABA is exploring too much and it intends to move the bats near to the best solution and decreases the step size. If the success-rate is more than 1/5, NABA is exploiting local optima too much and it increases the step size. That will help bats to reach the local optima faster, but the accuracy will decrease because of the long step size.

4. HILL CLIMBING ALGORITHM (HC)

Hill Climbing belongs to the family of local search and it is related to gradient ascent (Russell and *et al.*, 1995). Here, we start the search from a random position and iteratively test new candidate solutions around the current solution. If there is a better solution among the candidate solutions, adopt the new one. This technique improves the current value of the objective function until we reach a local optimum. When no local move could further improve, the search would have reached a local optimum of the objective function (Selman and Gomes 2006). The pseudo code of HC technique is shown in Algorithm 2.

Algorithm 2 Hill Climbing
1. S some initial candidate solution
2. **repeat**
3. R Tweak(Copy(S))
4. **If** Quality (R) > Quality(S) **then**
5. S R
6. **until** S is the ideal solution or we have run out of time
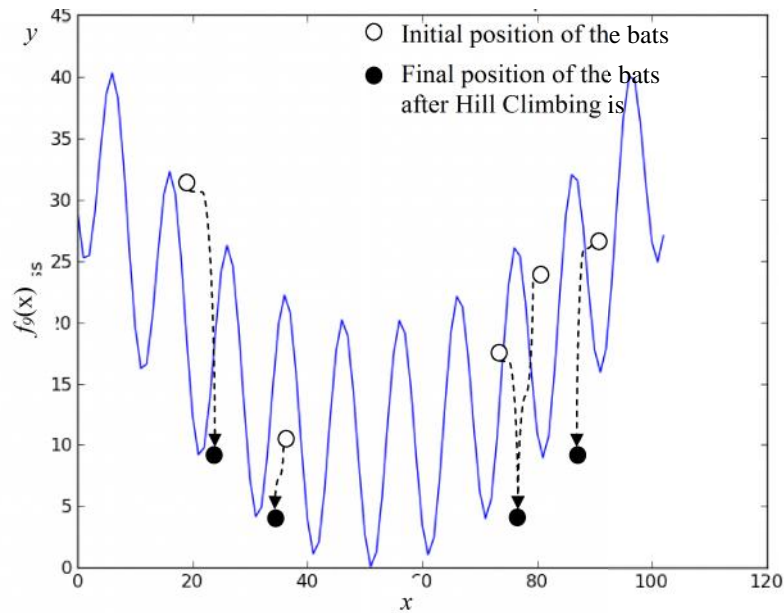7. **return** S

Fig. 1: The bats' positions before and after Hill Climbing on one dimensional Rastrigin's function.

5. PROPOSED HYBRID OPTIMIZATION ALGORITHM (BAHC)

Here, we propose a new Bat Algorithm (BA) which is hybridization of the Bat Algorithm and Hill Climbing, called BAHC. BA suffers from the premature convergence problem for high dimensional problems because the algorithm tends to converge very fast.

In BAHC, HC technique is used in two parts of BA. At the beginning of BA, the population is randomly initialized. Once initialized, HC is used to move all the bats to their local best. So, many local optimum traps will be detected at the beginning of the search.

Figure 1 depicts bats movement on one dimensional Rastrigin function when Hill Climbing algorithm is used. The y-axis shows the fitness values of Rastrigin function, x-axis shows the independent variables, and small circles on the function represent the bats solutions.

As it could be noted, at the beginning of the algorithm, Hill Climbing helps bats to move to the bottom of the basin (the algorithm is used to find the minimum value/fitness of Rastigin function) with a subsequent start of the bat algorithm.

In addition, whenever there is an improvement on bat's solution, Hill Climbing could enhance the improvement. So, bats will keep moving to their local optimum during the search with highest probability of locating the global optimum. The

BAHC pseudo-code is illustrated in Algorithm 1. The differences from BA are shown in boldface font.

Algorithm 1 Hybrid of Bat Algorithm with Hill Climbing - BAHC pseudo code.

1. *Objective function f(x), x=(x_1, ... , x_d)^T*
2. *Initialize the bat population x_i, v_i (i=1,2,...,n)*
3. **Modify the solutions by using Hill Climbing**
4. *Define pulse frequency f_i at x_i*
5. *Initialize pulse rates r_i and the loudness A_i*
6. while(t<*Max number of iterations*)
7. *Generate new solutions by adjusting frequency,*
8. *and updating velocities and locations/solutions [equations (1) to (3)]*
9. if*(rand > r_i)*
10. *Select a solution among the best solutions*
11. *Generate a local solution around the selected best solution*
12. end if
13. *Generate a new solution by flying randomly*
14. if*(rand < A_i & $f(x_i) < f(x^*)$)*
15. **Modify the solution by using Hill Climbing**
16. *Accept the new solutions*
17. *Increase r_i and reduce A_i*
18. end if
19. *Rank the bats and find the current best x^**
20. end while
21. *Post process results and visualization*

As a result, BAHC differs from the original Bat Algorithm in line 3 and 15, where solution is improved using Hill Climbing technique.

6. NUMERICAL EXPERIMENTS AND RESULTS

6.1 Simulation

In the present investigation 11 test functions were carried out for a comprehensive comparison (Yao *et al*., 1999). The benchmark functions f_1 - f_5 are unimodal functions. The functions f_8 - f_{11} are multimodal functions where the number of local minima increases exponentially with the problem dimension (Yao *et al*., 1999) that makes these functions are very hard to optimize. The test functions are described in Table 1.

BAHC was compared with original BA, CLSB, ABA and NABA to demonstrate the effectiveness of BAHC. In our experimental studies, parameter settings of the algorithms are the same as in their original papers. The population size of the algorithms has been kept as 50 regardless of the dimension size of the problem.

Table 1. Descriptions of the test functions

No	Name	Formula	Global Optimum	Search Domain
f_1	Sphere Model	$f_1(x) = \sum_{i=1}^{D} x_i^2$	0	$-100 \leq x_i \leq 100$
f_2	Schwefel's Problem 2.22	$f_2(x) = \sum_{i=1}^{D} |x_i| + \prod_{i=1}^{D} |x_i|$	0	$-10 \leq x_i \leq 10$
f_3	Schwefel's Problem 1.2	$f_3(x) = \sum_{i=1}^{D} (\sum_{j=1}^{i} x_j)^2$	0	$-100 \leq x_i \leq 100$
f_4	Schwefel's Problem 2.21	$f_4(x) = max_i\{|x_i|, 1 \leq i \leq 30\}$	0	$-100 \leq x_i \leq 100$
f_5	Generalized Rosenbrock's Function	$f_5(x) = \sum_{i=1}^{D}[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	0	$-30 \leq x_i \leq 30$
f_6	Step Function	$f_6(x) = \sum_{i=1}^{D}(\lfloor x_i + 0.5 \rfloor)^2$	0	$-100 \leq x_i \leq 100$
f_7	Quartic Function i.e. Noise	$f_7(x) = \sum_{i=1}^{D} x_i^4 + random\ [0,1)$	0	$-1.28 \leq x_i \leq 1.28$
f_8	Generalized Schewefel's Function 2.26	$f_8(x) = 12569.5 - \sum_{i=1}^{D} (x_i \sin\ (sqrt(|x_i|))$	0	$-500 \leq x_i \leq 500$
f_9	Generalized Rastrigin's Function	$f_9(x) = \sum_{i=1}^{D}[x_i^2 - 10\cos(2\pi x_i) + 10]$	0	$-5.12 \leq x_i \leq 5.12$
f_{10}	Ackley's Function	$f_{10}(x) = -20exp\ (-0.2\ (1/30\sum_{i=1}^{D} x_i^2)^{1/2}) - exp(1/30 \sum_{i=1}^{D} \cos(2\pi x_i)) + 20 + e$	0	$-32 \leq x_i \leq 32$
f_{11}	Griewank Function	$f_{11}(x) = \sum_{i=1}^{D} x_i^2/4000 - \prod_{i=1}^{D} \cos(x_i/sqrt(i)) + 1$	0	$-600 \leq x_i \leq 600$

Maximum number of function evaluation (FES) is set to 100.000 and the dimensions of the problems 30-D and 10-D for all the problems.

Figure 2 depicts the convergence characteristics in terms of the best fitness value of the median run of each algorithm, which is plotted using a logarithmic scale to reduce the biggest and smallest values in the whole optimization process. Some of the illustrations have been omitted in order to save space.

6.2 Simulation Results and Performance Assessment

To evaluate the performance of the algorithms each test functions were carried out 20 independent runs by the algorithms. After these runs the mean of the results and the standard deviation (STDEV) are recorded for each algorithms and test functions which are reported in Table 2. The best results are typed in bold.

In Table 2, the proposed BAHC has demonstrated a better ability of global searching for all the benchmark functions except for function f_4 and f_7. BAHC

achieves notable improvement on functions f_1, f_3, f_5, f_6, and f_8. Although BAHC outperformed other algorithms for the function f_8, it appears that all the algorithms trapped into similar local optima. It can be seen from Table 2 that using Hill Climbing with Bat Algorithm significantly increased the precision of the optimization results.

For the function f_7, BAHC performed very poorly comparing with NABA, CLSBA, and BA. f_7 is Quartic Function with noise which has ill-behaved surface. The locations of its minima and maxima are very close, so using Hill Climbing cost too much time to BAHC.

BAHC showed better performance than other algorithms on f_2, f_{10}, and f_{11} when the functions are 30 dimensions, but it didn't show same performance when the functions are 10 dimensions. Because BAHC loses too much time to reach the local minima at the beginning of the algorithm.

Furthermore, Fig.2 shows the convergence process of NABA, ABA, CLSBA, BA, and BAHC, conducted on some of the test functions from all the groups. The simulation results show the fast converging ability of BAHC in comparison with BA for most of the benchmark functions. Especially if one of the bat is on basin which contains global optima, the Hill Climbing part of BAHC increases the convergence speed rapidly as it can be seen for the function f_1, f_8, f_{10}, and f_{11}.

A close inspection of Table 2 reveals that, out of 11 functions, BAHC outperformed all other contestant algorithms on 9 functions in terms of Mean. It can be clearly seen that BAHC can provide a better optimization solution for most of the functions encompassing different type of functions with high dimensions. Therefore, it can be said that BAHC has the best universality on different type of problems. Additionally, we can say that, NABA has the best performance in terms of the mean after BAHC.

6.3 Algorithm Complexity

This section describes the algorithm complexity of BA and BAHC as defined in (Liang *et al.*, 2013). All the algorithms are developed in the Python environment and run on a PC with a 3.20 GHz CPU and 6.00 GB of RAM. For the random number generators, the Mersene Twister is used by Python in our simulations (Matsumoto and Nishimura 1998). Table 3 shows the BA and BAHC complexity on 10, 30, and 50 dimensions. As defined in (*Liang et al.*, 2013), T0 is the time calculated by running the following test problem:

```
for i=1:1000000
        x=0.55+ (double) i;
        x=x+x;
        x=x/2;
        x=x*x;
```

```
x=sqrt(x);
x=log(x);
x=exp (x);
x=x/(x+2);
end
```

T1 is the time to execute 200,000 evaluations of benchmark function f_9 by itself with D dimension, and T2 is the time to execute algorithms with 200,000 evaluations of f_9 in D dimensions. T2' is the mean T2 values of 5 runs. Based on Table 3, both T1 and T2' scaled linearly with the number of dimensions, as shown by the linear growth of (T2'-T1)/T0 for both algorithms. However, BAHC stayed behind the BA in all dimensions. Hill Climbing is slowing down BAHC.

Table 2. Performance comparison of Means and Standard Deviations for NABA, ABA, CLSBA, BA, and BAHC

f	D	NABA Mean (StdDev)	ABA Mean (StdDev)	CLSBA Mean (StdDev)	BA Mean (StdDev)	BAHC Mean (StdDev)
f_1	10	0.0053(0.0019)	0.1493(0.0245)	0.0068(0.0002)	0.0076(0.0014)	**0.0020(0.0005)**
	30	0.1216(0.0073)	1.7100(0.2421)	0.1238(0.0018)	0.1096(0.0104)	**0.0258(9.2E-04)**
f_2	10	**0.0784(0.0103)**	0.9560(0.1061)	0.0859(0.0069)	0.0795(0.0056)	0.0828(0.0116)
	30	0.5895(0.0281)	6.5875(0.0252)	0.6274(0.0076)	0.6570(0.0237)	**0.5372(0.0461)**
f_3	10	7.3106(2.2131)	10.3595(4.7663)	8.4325(1.2690)	7.5439(1.415)	**4.5081(1.0634)**
	30	628.78(62.36)	786.59(85.43)	732.17(98.45)	620.45(74.34)	**227.93(20.74)**
f_4	10	**0.0814(0.0075)**	0.2065(0.0240)	0.0979(0.0153)	0.0874(0.0056)	0.0990(0.0146)
	30	0.3193(0.0256)	0.5905(0.0652)	0.3185(0.0431)	**0.2924(0.0013)**	2.3131(0.5116)
f_5	10	63.36(68.99)	29.09(13.06)	12.26(1.111)	73.10(121.11)	**8.844(0.874)**
	30	223.61(116.29)	335.94(14.90)	186.40(21.63)	350.39(189.86)	**54.07(5.67)**
f_6	10	5.149(0.568)	8.546(0.853)	15.964(2.541)	12.347(0.982)	**2.154(0.0254)**
	30	48.69(5.8083)	58.12(6.2412)	52.82(8.1370)	54.09(1.4475)	**10.10(0.8899)**
f_7	10	**0.0005(0.0002)**	0.0817(0.0231)	0.0007(0.0005)	0.0009(0.0003)	0.0011(0.0008)
	30	**0.0045(0.0014)**	5.8277(0.7727)	0.0047(0.0010)	0.0086(0.0028)	1.0357(0.5603)
f_8	10	1405.62(265.21)	1229.85(347.98)	1686.09(389.28)	1509.57(467.18)	**1018.62(159.54)**
	30	5274.23(436.74)	5621.54(379.73)	5549.87(1275.91)	6384.07(376.02)	**4158.53(185.24)**
f_9	10	47.37(20.89)	54.79(11.76)	53.79(17.52)	39.79(7.233)	**38.87(6.466)**
	30	196.76(35.25)	317.21(35.27)	214.28(8.80)	246.92(11.27)	**172.61(20.22)**
f_{10}	10	8.322(8.430)	**5.447(7.257)**	15.52(6.904)	8.958(8.975)	15.34(6.827)
	30	19.81(0.2147)	19.96(8.1E-04)	20.00(0.2873)	19.94(0.0161)	**19.72(0.1157)**
f_{11}	10	**0.6892(0.0792)**	0.8576(0.0565)	0.8168(0.0308)	0.7886(0.0581)	0.7518(0.0665)
	30	1.4284(0.0135)	1.4794(0.0979)	1.5004(0.0147)	1.4693(0.0067)	**1.0791(0.0101)**

Table 3. Algorithms complexity

BA	T0	T1	T2`	(T2`-T1)/T0
D=10	8.692	15.184	15.553	0.042
D=30		16.792	17.329	0.0618
D=50		17.580	18.834	0.086
BAHC	T0	T1	T2`	(T2`-T1)/T0
D=10	8.692	16.779	16.824	0.005
D=30		16.851	17.411	0.064
D=50		17.693	18.988	0.148
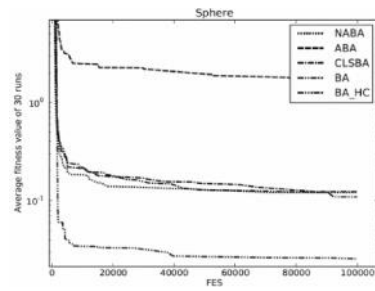
f_1-30D
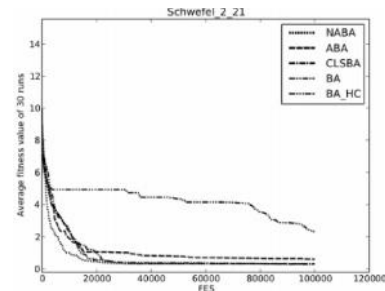
f_4-30D
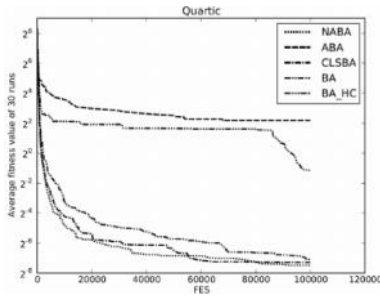
f_7-30D
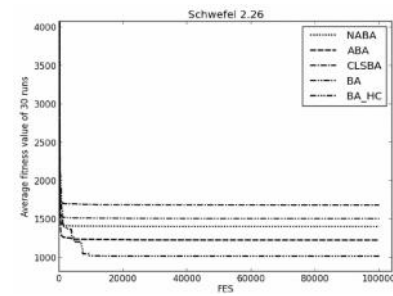
f_8-10D
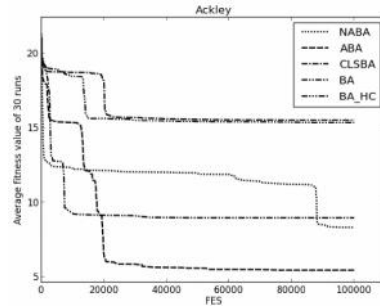
f_{10} -10D
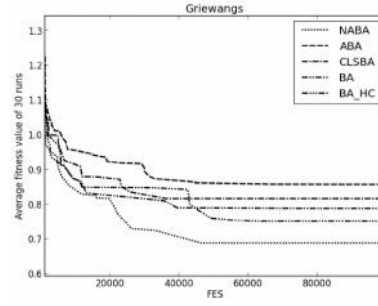
f_{11} -10D

Fig. 2: Convergence characteristics of NABA, ABA, CLSBA, BA, and BAHC on the functions $f_1, f_4, f_7, f_8, f_{10}$, and f_{11}.

7. CONCLUSIONS AND FUTURE WORKS

The present paper provides some information about BAHC, a hybrid of bat algorithm and hill climbing to solve high dimensional function optimization problems. Its performance was compared with those of NABA, ABA, CLSBA, and BA over 11 optimization problems. Results reported that the mixing BA with HC enhanced the optimization accuracy and improved the computation time of BA. Evaluation results obtained on optimization functions proved the effectiveness of the hybrid algorithm.

Although the performance of BAHC outperforms BA and other algorithms in terms of accuracy and computation time for high-dimensional problems, BAHC does not perform so effectively for low dimensional problems which need low computation load. However BAHC spends too much time during hill climbing and then starts converging. As a future work convergence ability of BAHC for low-dimensional optimization problems would be studied.

REFERENCES:

Baziar A, Kavoosi-Fard A, Zare J. 2013. A novel self-adaptive modification approach based on bat algorithm for optimal management of renewable MG. *Journal of Intelligent Learning Systems and Applications*, **5(01)**; 11-18.

Chandrasekar C. 2013. An optimized approach of modified BAT algorithm to record deduplication. *International Journal of Computer Applications*, **62(1).**

Dorigo M, Maniezzo V, Colorni, A. 1996. Ant system: optimization by a colony of cooperating agents. *Systems, man, and cybernetics, Part B: Cybernetics, IEEE Transactions*, **26(1)**; 29-41.

Gandomi AH, Yang XS, Alavi AH, Talatahari S. 2013. Bat algorithm for constrained optimization tasks. *Neural Computing and Applications*, **22(6)**; 1239-1255.

Kabir MWU, Sakib N, Chowdhury SMR, Alam MS. 2014. A Novel Adaptive Bat Algorithm to Control Explorations and Exploitations for Continuous Optimization Problems. *International Journal of Computer Applications*, **94(13) 15-20**.

Karaboga D. 2010. Artificial bee colony algorithm. Scholarpedia **5 (3)**: 6915.Author's Photo.

Kennedy J, Eberhart RC. 1995. Proceedings of IEEE International Conference on Neural Networks.Perth, Australia.

Kim J, Bentley PJ, Aickelin U, Greensmith J, Tedesco G, Twycross

J. 2007. Immune system approaches to intrusion detection–a review. *Natural computing,* **6(4)**; 413-466.

Liang JJ, Qu BY, Suganthan PN, Hernández-Díaz AG. 2013. Problem definitions and evaluation criteria for the CEC 2013 special session on real-parameter optimization. *Computational Intelligence Laboratory*, Zhengzhou University, Zhengzhou, China and Nanyang Technological University, Singapore, Technical Report, 01212.

Lim A, Lin J, Rodrigues B, Xiao F. 2006. Ant colony optimization with hill climbing for the bandwidth minimization problem. *Applied Soft Computing*, **6(2)**, 180-188.

Matsumoto M, Nishimura T. 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, **8(1)**.3-30.

Rechenberg I. 1994. Evolutionsstrategie'94, volume 1 of Werkstatt Bionik und Evolutionstechnik.Frommann {Holzboog, Stuttgart.

Renders JM, Bersini H. 1994. Hybridizing genetic algorithms with hill-climbing methods for global optimization: two possible ways. *In Evolutionary Computation, 1994*. IEEE World Congress on Computational Intelligence. *Proceedings of the First IEEE Conference,* 312-317. IEEE.

Richardson P. 2008. The Secret Life of Bats. [Online] http://www.nhm.ac.uk/nature-online/life/mammals/bats/session3/

Russell S, Norvig, P. Intelligence A. 1995. A modern approach. Artificial Intelligence. Prentice-Hall, Englewood Cliffs, **25**; 111-113.

Selman B, Gomes CP. **2006**. Hill-climbing search. Encyclopedia of cognitive Science.

Topal AO, Altun, O. 2014. Dynamic Virtual Bats Algorithm (DVBA) for Global Numerical Optimization. International Conference on *Intelligent Networking and Collaborative Systems (INCoS);*320-327. IEEE.

Wang X, Wang, W, Wang Y. 2013. An adaptive bat algorithm. *In intelligent computing theories and technology*, 216-223. Springer Berlin Heidelberg.

Yang XS. 2009. Firefly algorithms for multimodal optimization. *Stochastic algorithms: foundations and applications*, 169-178. Springer Berlin Heidelberg.

Yang XS. 2010. A new metaheuristic bat-inspired algorithm. *Nature inspired cooperative strategies for optimization (NICSO 2010);* 65-74. Springer Berlin Heidelberg.

Yang XS. 2011. Bat algorithm for multi-objective optimization. *International Journal of Bio-Inspired Computation*, **3(5)**; 267-274.

Yang XS, Deb S. 2009. Cuckoo search via Lévy flights. *In nature & biologically inspired computing*, World Congress on NaBIC 2009. 210-214. IEEE.

Yao X, Liu Y, Lin G. 1999. Evolutionary programming made faster. IEEE Transactions on Evolutionary Computation, **3(2)**, 82-102.

`Yildiz AR. 2009.` An effective hybrid immune-hill climbing optimization approach for solving design and manufacturing optimization problems in industry. *Journal of Materials Processing Technology*, **209(6)**; 2773-2780.

Yin PY, Yu SS, Wang PP, Wang YT. 2006. A hybrid particle swarm optimization algorithm for optimal task assignment in distributed systems. *Computer Standards & Interfaces*, **28(4)**, 441-450.

Yuanbin M, Xinquan Z, Shujian X. 2013. Local memory search Bat algorithm for grey economic dynamic system. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, **11(9),** 4925-4934.

REST ARCHITECTURE STATE OF PRACTICE IN MACEDONIAN IT COMPANIES

Agon MEMETI
Department of Informatics, Faculty of Math and Natural Sciences,
University of Tetovo, Macedonia
Florinda IMERI
Department of Informatics, Faculty of Math and Natural Sciences,
University of Tetovo, Macedonia
Betim ÇIÇO
Faculty of Contemporary Siences and Technologies,
South East European University

ABSTRACT

Rebuilding applications by reusing the existing ones is of primary importance for the IT developers. Every system uses resources such as Web pages, business information, etc. that can be represented in a computer-based system because the main purpose of IT department is making these resources accessible to its clients. Service architects and developers are interested in creating services that are implementable, maintainable, extensible and scalable. It is the RESTful design that promises this and a lot more. It has been more than a decade since the introduction of *Representational State Transfer* (REST), and it seems to become one of the most important technologies for Web applications. Every major development language now includes frameworks for building RESTful Web services. Considering the importance of REST while reusing services, we have conducted a survey in IT companies in Macedonia, which relies upon collection of empirical data. The survey aimed at exploring developers' experience on REST Architecture in IT Companies and the data are here reported and evaluated.
Keywords: REST Architecture, service reuse, resources, empirical data.

1. INTRODUCTION

The products or services offered by businesses today rely on information products developed by enterprise IT department. Time to market of a product is very significant for enterprises in order to be competitive in the market. As a result, minimizing the production time by IT departments for a given product is essential for the enterprises (Algermissen, 2011).

The primary attraction for developers is to rebuild applications by reusing the existing ones. Reusing existing objects/components or making use of class libraries, functions, data, etc is as old as retelling a story (Imeri and Antovski, 2012). To avoid duplicate designs of similar processes, abstract business processes were used to support for reusability.

Every system uses resources, such as Web pages, business information, or anything that can be represented in a computer-based system. The purpose of a service is to provide to its clients an easy access to these resources. Also service architects and developers are interested in creating services that are easy to implement, maintain and are extensible, and scalable. It is the RESTful design that promises this and a lot more.

It has been more than a decade since REST introduction, and it seem to become one of the most important technologies for Web applications. Its importance seems to grow very quickly since all technologies are moving towards Application Programming Interface (API) orientation. Every major development language now includes frameworks for building RESTful Web services. As such, it is important for Web developers and architects to have a clear understanding of *Representational State Transfer* (REST) and RESTful services. Many books and tools have been created but there is still a general lack of understanding its fundamentals as an architecture style. The reason perhaps could be found in the fact that REST was presented in a doctoral dissertation, with relatively high entry barriers for its understanding, or because the description used models that were more oriented towards documentation than to working practitioners (Wilde and Pautasso, 2011).

Seeing the importance of REST while reusing services, we have conducted a survey in IT companies in Macedonia, which relies upon collection of empirical data. The survey aimed to explore developers' experience on REST Architecture, its benefits and challenges from their viewpoint and see the possible contributors towards successful reuse, which could help increase of the knowledge and understanding of REST. We have prepared a questionnaire and sent it electronically to developers in 30 companies in Macedonia and other Balkan countries, but only the companies in Macedonia replayed to the questionnaire. Consequently, comparing the results obtained from IT companies in Macedonia and the region failed.

The forthcoming part of the paper reports on the state of the art of the REST Architecture, the questionnaire and research questions and empirical results. In the end, the conclusions are drawn.

2. REST ARCHITECTURE

Restful Services as architectural style lighter than SOAP-based Web Services, due to its simplicity, heterogeneity and web-based format, entities/resources are identified by unique URLs, including how resource states are addressed and transferred over HTTP by a wide range of clients written in different languages (Memeti *et. al.*, 2015).

REST originated at the intersection of academia and software development, among the architects of the World Wide Web (Adamczyk *et al.*, 2011). Unfortunately, researchers have only recently started to work on RESTful services. No paper about RESTful Web services has been reported up to 2007. Information about RESTful services was reported only in 2010.

The REST architectural style has been chosen for the resource centric approach because it could be easily applied and naturally mapped to it. The HTTP methods can be mapped to the Create (POST), Retrieve (GET), Update (PUT) and Delete (DELETE) operations and the XML documents can be used to provide a uniform resource representation (Szepielak, 2006).

Many frameworks and tools for building RESTful Web services are available today. They are written in different programming languages and range from simple to quite sophisticated in their support of HTTP and other Web technologies. As they continue to improve, misunderstandings and violations present in today's Web services will likely lessen (Wilde and Puatasso, 2011).

When multiple providers offer the same service, a client has a choice and can select the most suitable one. Often, this choice comes down to the Quality of Service (QoS) parameters (Wilde, 2011). RESTful Web services today ignore QoS requirements; their only concern is providing functional interfaces. To add QoS parameters to RESTful services, a language for describing the parameters and a mechanism to incorporate the description in the HTTP payload is needed. Defining a standard QoS description language might benefit from the work in Semantic Web. Semantic Web ontologies define standard ways of interpreting information, such as QoS parameters, enabling all clients to interpret them the same way (Wilde, 2011).

Building an integrated software environment in an enterprise often requires developing large amounts of Web services. The integration efforts can be greatly reduced by using a specialized framework for their development. Providing such tools that simplify software development in integration projects is essential for optimizing their efficiency and cost. One of the most important choices to make when building an integration solution is to select an appropriate integration approach and suitable technologies for its realization (Wilde, 2011).

According to Fielding, (2000) the four principles of REST (called constraints) are: i) identification of resources, ii) manipulation of resources through

representations, iii) self-descriptive messages and, iv) hypermedia as the engine of application state.

These principles describe the architecture of systems and interactions that make up the Web. The building blocks of the Web are called resources. A resource is anything that can be named as a target of hypertext (e.g., a file, a script, a collection of resources). In response to a request for a resource, the client receives a representation of that resource, which may have a different format than the resource owned by the server. Resources are manipulated via messages that have standard meanings; on the Web, these messages are the HTTP methods (Fielding, 2000).

Activities that companies are supporting such as knowledge management and communication or decision making, the development approach such as object -oriented or component-based approach which is characterized by reusability (elements for re-used in other workflows), substitutability (alternative implementations easily inserted with precisely specified interfaces, and with the ability to verify and validate substitutions), extensibility and scalability (the ability to extend system component and to scale it, by increasing capabilities of individual components, with new functionalities), customizability (ability to add new features based on the needs of a particular domain), and composability (easy construction of complex solutions using basic components and the ability to estimate the efforts of system integration and compose) are of great importance for the developers.

According to (Choi and Kim, 2008) the advantages of REST architecture are: i) *availability:* utilization of service which increases the reusability service, ii) g*enerality:* a more generic service, since a reusable service could be implementable for unknown future requirements, iii) *understandability:* service and its interface are understandable. The more understandable it is, the more it is used, iv) *functionality:* to have a highest degree of guaranteed reuse, the service should pack with a complete range of functionality, v) *reliability:* since the service is to be used by many consumers, it should be reliable, be able to operate without any failure continuously and, vi) *portability:* the portable services which are to be consumed through any kind of environment without being modified increases the reusability service.

3. EMPIRICAL ANALYSIS

The present paper reports on usability of the REST architecture, its advantages and challenges. Consequently, a questionnaire with predefined answers and some open questions for general reflection was compiled. The research involved both qualitative and quantitative studies. In order to have clear understanding, at some points, some of the data were transformed from qualitative into

quantitative presentations.

As regarding the sampling frame and size of our survey, the data presented in this paper are answers taken only from 10 software developing companies in Macedonia. The questionnaire was electronically sent to 20 companies in the region, but 10 companies in Macedonia replayed to the questions, only.

Reported in Table 1, the questionnaire consists of 13 questions which have not been made before. Below are presented research questions of our survey.

1. RQ1. Do companies use REST Architecture?

2. RQ2. Does developers' experience with REST influence to the percentage of service reuse?

3. RQ3. Which factors influence toward use of REST Architecture?

4. RQ4. Which challenges and benefits are reported from companies?

Table 1. The questionnaire

Q1.Which Architectural Styles and Patterns are used by your organization?
Q2.What is your experience with REST Architecture?
Q3.Which activities are supporting the architectural styles used by your company?
Q4.Which development approach is used by your organization?
Q5.When a development team set up a project, would they start to go through old services or just start to write the new program application?
Q6.What is the percentage of service reuse?
Q7.How is the software production process carried out in your organization?
Q8.While enhancing performances of existing applications, do you have any problems with permissions to use same data?
Q9.If yes, how do you resolve the problem?
Q10.Does the quality and flexibility of applications/services increased using REST (based on your experience)?
Q11.Does the productivity of applications/services increases using REST (based on your experience)?
Q12.Which advantages of REST Architecture are more important from your point of view?
Q13.Which challenges do you face while using REST Architecture?

3.1 Results

The survey involved software development companies, with more than 3 years working experience, whose application domain is information systems such as administrative, commerce and educational software and the data of the survey are in forthcoming part of the paper provided.

A) RQ1. Do companies use REST Architecture?

Answers to questions Q1 gave us information about the way companies develop software, i.e. which architectural style are using in their company. From the answers we could see that companies use different styles. Some have chosen one and some have chosen more than one style. The results were summed up and presented as percentage data to find the mostly used architecture. Depicted in the Figure 1, the results reported that 60% of them use REST architecture; 30% component-based applications; 20% data-centric; 10% event/driven; 40% service-oriented and 60% client/server which gives the same percentage like REST usage.
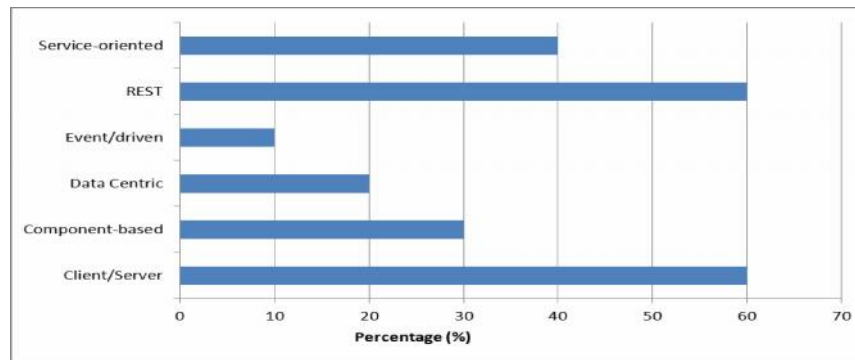

Fig. 1: Percentage Architectural Styles used by Software Companies.

B) RQ2. Does developers' experience with REST influence to the percentage of service reuse?

Since REST is relatively a new approach, our second research questions tries to see the relations among developers experience with REST and the percentage of service reuse. Based on surveys on software reuse (Imeri and Antovski, 2011) which is closely related to component based software reuse, we hypothetically concluded that the more experienced developers are, the higher the reuse is. The results obtained from the question Q2 reported that 80% of the developers had high experience with REST. 20% of the developers had low experience. The results on reuse service obtained from question Q6 reported different values (Figure 2). So, the developers' experience with REST merely has a considerable impact on service reuse.
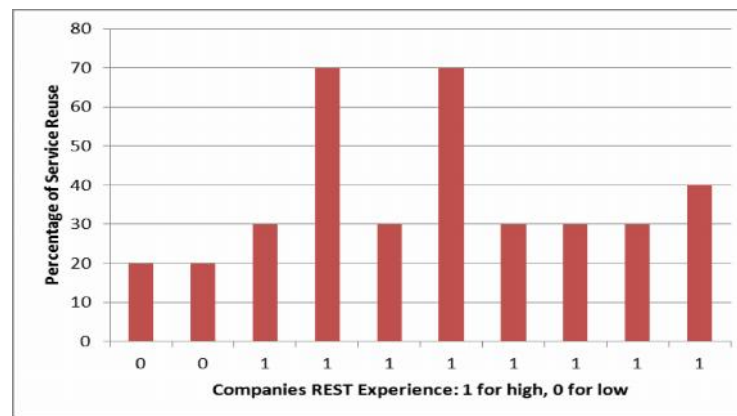
Fig. 2: REST experience vs. service reuse

C) RQ3. Which factors influence toward use of REST Architecture?

Information about the factors influencing the use of REST Architecture is from the questions Q3, Q4, Q5 and Q7 provided information.

80% of companies are using Knowledge Management and Communication (KMC), and 20% of them Design Reasoning and Decision Making (DRDM).

Regarding the development approach, 70% of companies have chosen object oriented (OO) paradigm and 30% of companies have chosen component-oriented (CO) paradigm. When a development team set up a new project, 60% of companies responded that while they develop new applications they go through old services (OS) whereas 40% start to write new program (NP).

Finally, questions regarding the software production process, 50% of responders answered that the organization produce specific software (isolated products - IO) whereas 50% product families (PF), meaning that organization develops product that evolves over time, being somehow adopted for each new project (in-house products).

Contingency tables (Table 2) were used to investigate the influence/dependence of the aforementioned factors by displaying the contingency distribution of variables. As we had 2x2 tables for the Chi-squared test, the Yates' continuity correction was applied.

We have applying the Yates' Chi Squared Test (as one of statistical test applied to sets of categorical data), since the effect of Yates' correction is to prevent overestimation of statistical significance for small data (https://en.wikipedia.org/wiki/Yates%27s_correction_for_continuity, 2015), based on the fact that the number of our responders is small.

Table 2 reports the Contingency tables of observed and expected frequencies of different variables which influenced the use of REST Architecture such as: i) the interrelation between the uses of REST Architecture vs. Companies development approach, ii) the use of REST Architecture vs. Activities supporting the architectural style, iii) the use of REST Architecture vs. Software production process and, iv) the use of REST Architecture vs. Developing process. Lastly, the significance of the difference between the two variables is assessed with a variety of statistical tests including Yates' chi-squared test.

The critical value of $S«^2$ at a 5% significance level for 1 degree of freedom is seen to be 3,841 (https://en.wikipedia.org/wiki/Chi-squared_distribution). Below, are presented the calculated values of $S«^2$. Almost in all our cases taken as factors that influence the use of REST Architecture, the calculated $S«^2$ value were smaller than the critical value of $S«^2$, i.e., independent variables except the software production process found near the critical value of $S«^2$, and can be taken as a dependent variable, such a factor that influence the use of REST Architecture.

Table2. Contingency tables of factors influencing the use of REST.

		CO	OO	Total				KMC	DRDM	Total
	REST	2	4	6			REST	5	1	6
Observed	Others	1	3	4		Observed	Others	2	2	4
	Total	3	7	10			Total	7	3	10
		CO	OO	Total				KMC	DRDM	Total
Expected	REST	1,8	4,2	6		Expected	REST	4,2	1,8	6
	Others	1,2	2,8	4			Others	2,8	1,2	4
	Total	3	7	10			Total	7	3	10
x^2 Yetes'	0,607143					x^2 Yetes'	1,638889			
		OS	NP	Total				IO	PF	Total
	REST	3	3	6			REST	4	2	6
Observed	Others	3	1	4		Observed	Others	1	3	4
	Total	6	4	10			Total	5	5	10
		OS	NP	Total				IO	PF	Total
Expected	REST	3,6	2,4	6		Expected	REST	3	3	6
	Others	2,4	1,6	4			Others	2	2	4
	Total	6	4	10			Total	5	5	10
x^2 Yetes'	1,100694					x^2 Yetes'	3,5			

D) RQ4. Which challenges and benefits are reported from companies?

The challenges of REST consist of data permissions. Results reported that 40% of developers have had problems with data permission due to their sensibility. 60 % of developers have not had problems with permission. Results are in Figure 3 depicted.
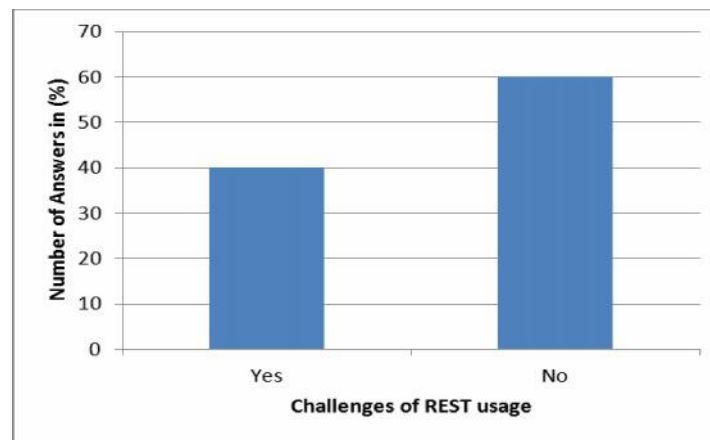
Fig. 3: Data permissions problems.

Figure 4 depicts some of the characteristics of the REST architecture. 80% of the companies reported that they use the Rest architecture because of its flexibility. 60% of the companies reported that they use the Rest architecture because of its scalability.
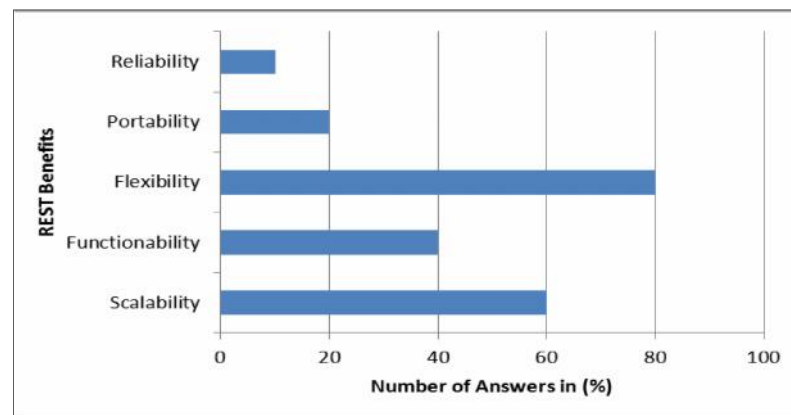

Fig.4: REST Architecture benefits.

4. CONCLUSIONS

The presented investigation addresses the companies' efficiency when meeting the customers' requests and demands for new products. In general the empirical data collected represent an overview of the current situation among companies using the REST architecture. The most evident patterns that could be identified

from our analysis, such factors that influence toward the use of REST Architecture are presented in Table 2, where variable dependencies are calculated.

The data statically (empirically) reported that the use of REST Architecture in Macedonian companies is relatively high, about 60%. Developers' experience merely has any impact on service reuse.

The factors that influence the use of REST architecture are Table 2 reported, showing that low probabilities of the tested variables dependence among factors we have chosen. A Yetes' value ($S«^2= 1, 638889$) is the calculated value among the use of REST architecture and activities supporting the architectural styles. As far as the development approach was concern, the calculated variables show that ($S«^2= 0, 607143$) is the Yetes' value among the development approach and the use of REST Architecture. In addition, a value ($S«^2= 1, 100694$) is reported as a Yetes' value while calculating the development approach such working through old services or write new program application as a activity supported by investigated companies, and lastly a Yetes' value ($S«^2= 3, 5$) based on the software production process. Consequently, all chosen variables are independent from the use of REST architecture. The process of software production taken as a variable in our case influences the use of REST because of the low probability of dependence, considering that the calculated $S«^2$ value is near the critical value of $S«^2$.

The advantages of REST consist of flexibility and productivity. Results showed that about 80% of applications that involved REST architecture showed positive impact, which is in line with the principles of Fielding (2000).

ACKNOWLEDGEMENTS

REFERENCES

Adamczyk P, Smith PH, Johnson RE, Hafiz M. 2011. REST and Web Services/ : In Theory and in Practice. REST: From Research to Practice. Spreinger. 35-57.

Algermissen J. 2011. Quantifying Integration Architectures. REST: From Research to Practice. Springer. 137–160.

Chi-squared distribution. Available online: https://en.wikipedia.org/wiki/Chi-squared_distribution.

Choi SW, Kim SD. 2008. A quality model for evaluating reusability of services in soa. E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing. E-Commerce and E-Services. ISBN: 978-0-7695-3340-7. 293-298.

Fielding R. 2000. Architectural Styles and the Design of Network-based Software Architectures. PhD Dissertation. 1-180.

Imeri F, Antovski L. 2012. An Analytical View on the Software Reuse. ICT Innovations 2012. Web Proceedings ISSN 1857-7288, 213.

Imeri F, Antovski L. 2013. An Empirical Study on Software Reuse in Small IT Companies in the Balkan Region**.** International Journal of Computers and Technologies. vol. 9, no. 2.1040–1048.

Memeti A, Selimi B, Besimi A, Cico B. 2015. A Framework for Flexible REST Services: Decoupling Authorization for Reduced Service Dependency. 4th Mediterranean Conference on Embeded Computing. IEEE Press. ISBN:978-9-9409-4364-6. 51-55.

Szepielak D. 2006. REST-Based Service Oriented Architecture for Dynamically Integrated Information Systems. PhD Symposium at ICSOC. 8–12.

Wilde E, Pautasso C. 2011. The Essence of REST Architectural Style. REST: From Research to Practice. Springer. 21–33.

Yates's correction for continuity. Available online: https://en.wikipedia.org/wiki/Yates%27s_correction_for_continuity.

MECHANICAL AND HYDRAULIC BEHAVIOUR OF GEOTEXTILES

Ilda KOLA and **Blerina KOLGJINI**
Department of Textile and Fashion, Faculty of Mechanics, Polytechnic University of Tirana, Albania

ABSTRACT

The present paper provides information about the properties of geotextiles based on their application area. Geotextiles are planar structures consisting of synthetic or natural polymeric material, used in contact with soil and other materials. Geotextiles are widely used in the construction industry due to their physical mechanical and hydraulic properties and durability. The present paper reports on the most important mechanical and hydraulic properties of geotextile used in specific applications areas along with their respective test methods based on Albanian and European Standards.

Keywords: geotextile, properties, test methods, standards

1. INTRODUCTION

The properties characterizing the geotextiles make them fit for application in civil engineering such as filtration, drainage, reinforcement, separation, protection and fluid barrier or strength release purposes. Each application requires geotextiles performing one or more specific functions and meeting some specific criteria related to fibers' characteristics and fabric construction, referred to as the structure of the geotextile (Karcz and Holtz 1988).

In addition to endurance and durability, geotextile are mentioned for their physical, mechanical, hydraulic properties. For each property, geotextiles are submitted to the respective testing method. Geotextiles is the combination of the physical form of the polymer fibers, their textile construction and the polymer chemical characteristics. In addition to the aforementioned properties, geotextile testing might be characterized as index or performance as a general value from which a specific property could be qualitatively assessed and used for general characterization of a geotextile product could be provided. Moreover, these tests do not provide data that can be directly used for design purposes and are performed on the geotextile alone, or in-isolation. For instant assessment

geotextiles are required to be investigated along with associating material such as soil, etc. As a result, information about the expected behavior of a geotextile in an engineered system could be obtained (Holtz *et al.*, 1998).

Physical properties help characterize the geotextile in the as-received, manufactured condition. Data on the physical index could be obtained via index testing. Common physical properties are specific gravity, mass per unit area, thickness and stiffness.

Long-term performance or durability of geotextiles is also affected by geotextile degradation via different mechanisms including ultraviolet light (sunlight), chemical reactions with geotextile polymers, and/or thermal degradation. Degradation testing is important in determining the ultimate lifetime of a geotextile in an engineered system (Zornberg and Thompson, 2012).

Mechanical responses and filtration ability are the most important properties of geotextile as they are a means to address working efficiency (Karcz and Holtz 1988). The present paper provides information about these properties.

2. Mechanical properties

The mechanical response of a geotextile depends on the orientation and regularity of the fibres and the type of polymer from which it is made. As aforementioned said, the geotextile must perform work in a stressed environment and be able to resist damage in an arduous environment.

Usually the stressed environment is known and the textile is selected on the basis of some criteria to cope with the expected imposed stresses and its ability to absorb those stresses over the proposed lifetime of the structure without straining more than a predetermined amount.

Figure 1 compares the tensile behaviors of a range of geotextiles (Horrocks and Anand, 2000). On the other hand, damage can be caused on site when constructing (e.g. accidental tracking from vehicles) or in-situ, when using geotextile (e.g. punching through geotextiles by overlying angular stone).

Performance ability is based on the textile stiffness in tension and its ability to resist creep failure under any given load condition. The ability to resist damage is complex, being a function of the fibre's ability to resist rupture and the construction of the fabric, which determines how stresses may be concentrated and relieved (Horrocks and Anand, 2000). The weight or area density of the fabric is an indicator of mechanical performance only within specific groups of textiles, but not between one type of construction and another. For example, within the range of needle punched continuous filament polyester fabrics, weight will correlate with tensile stiffness, while a woven fabric with a given area density will be much stiffer than an equivalent weight needle punched structure. So, the construction controls the performance. Therefore, weight cannot be

used alone as a criterion in specifying textiles for civil engineering use, but in combination with other specified factors, weight can be a useful indication of the kind of product required for a particular purpose (Christopher and Holtz, 1985).
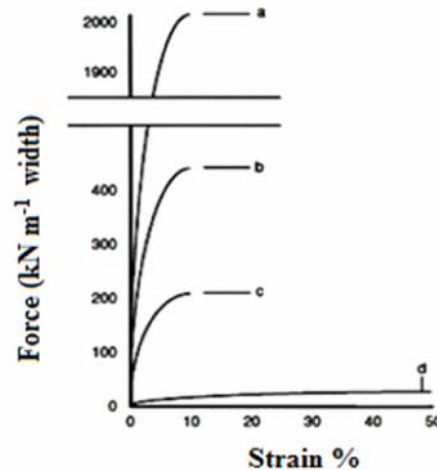

Fig.1: Typical ultimate stress–strain failure levels (a) of high strength and (b) of medium strength PES woven geotextiles used for embankment support and soil reinforcement, (c) of geogrids and lower strength polyester woven geotextiles used for soil reinforcement and (d) of low strength, highly extensible nonwoven geotextiles used for separation and filtration (Horrocks and Anand, 2000).

One of the most important strength properties for a geotextile is the *tensile strength*. The breaking strength of a standard width of fabric or 'ultimate strip tensile failure strength' is universally quoted in the manufacturers' literature to describe the 'strength' of their textiles. But this is of very limited use in terms of design. Rather, a strength at a given small strain level will be the design requirement. Therefore, the tensile resistance or modulus of the textile at say, 2%, 4%, and 6% strain is much more valuable. To understand the load-strain characteristic, it is important to consider the complete stress-strain curve. It is also important to consider the nature of the test and the testing environment (Holtz *et al*., 1998). Ideally, continuous stress–strain curves should be provided for engineers, to enable them to design stress resisting structures properly. Geotextile tensile strength is measured by clamping two opposite ends of a geotextile specimen in a mechanical testing machine and stretching the specimen until failure occurs. Typically, both the force applied to the geotextile and the geotextile strain are measured allowing observation of the stress-strain curve and development of associated modulus (Zornberg and Thompson, 2012).

Stress–strain curves, as shown in Fig. 1 and 2, may well comprise a high

strain sector, contributed by the textile structure straightening out, and a low strain sector, contributed by the straightened polymer taking the stress. In Figure 2, geogrids (a) absorb the imposed stresses immediately, giving a high initial modulus. Later, the curve flattens. Woven fabrics (b) exhibit initial straightening of warp fibres which produces a low initial modulus. Later the modulus increases as the straightened polymer fibres take the stress directly. Nonwovens (c) give a curvilinear curve, because extension is primarily resisted by straightening and realignment of the random fibre directions (Horrocks and Anand, 2000).

Geotextiles might have different strengths in different directions. Therefore, tests should be conducted in both principal directions. The wide-width tensile test S SH EN ISO 10319 is the most commonly specified in Albanian and European Standards for testing geotextiles used in almost all the applications performing all functions (Drejtoria e Përgjithshme e Standardizimit, 2008).
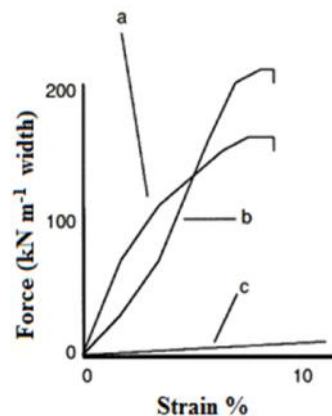

Fig.2: Different stress–strain curve shapes exhibited by the three main types of geosynthetic construction. (a) Geogrids, (b) Woven fabrics, (c) Nonwovens.

Tensile creep is a time-dependent mechanical property and strain at constant load. Creep can cause the physical failure of a geotextile if it is held under too high mechanical stress. Creep tests can be run for any of the tensile test types, but are most frequently performed on a wide strip specimen by applying a constant load for a sustained period. Short-term creep strain is strongly influenced by the geotextile structure. Woven geotextiles have the least; heat-bonded geotextiles have intermediate; and needled geotextiles have the most. Longer-term creep rates are controlled by structure and polymer type. The creep limit is the most important creep characteristic. It is the load per unit width above which the geotextile will creep to rupture. The creep limit is controlled by the polymer and for polyester it is approximately 60% of its ultimate strength, for polyethylene about 40% of its ultimate strength and for polypropylene around

20% of its ultimate strength. Therefore, for example, a polyester fabric with an ultimate tensile strength of 100kNm-1 width cannot be loaded under a long term stress of more than 60 kNm-1. The higher the level of imposed stress above this point, the more rapid will be the onset of creep failure. Figure 3 depicts the safe loading limits for most commonly used geotextiles (Horrocks and Anand, 2000). The S SH ISO 13431 '*Determination of tensile creep rupture behavior'* is defined in the Albanian Standards as a means to address tensile creep for a number of specific applications (Drejtoria e Përgjithshme e Standardizimit, 2008).
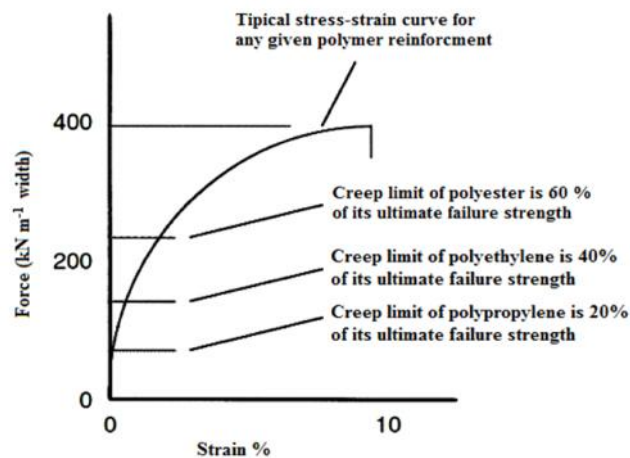

Fig. 3: Approximate limits of creep resistance for different geosynthetic polymer constructions (Horrocks and Anand. 2000).

Other tests for geotextile strength are focused on ensuring that a geotextile is strong enough to survive installation stresses, often the most severe stresses placed on the geotextile over its lifetime. The most commonly specified tests cover tear strength and puncture strength (Zornberg and Thompson, 2012).

Often, due to manufacturing size limitations, it is desirable to join multiple geotextile sections together when placed in an engineered system. The most common method of joining geotextiles is by sewing them. If sewing is used when installing geotextiles, characterization of the strength of a given seam type would be important for ensuring the integrity of the sewn sections. This is carried out by testing a tensile strength on a sewn geotextile sample as defined by the Albanian Standard, S SH EN ISO 10321 (Drejtoria e Përgjithshme e Standardizimit, 2008).

3. Hydraulic Properties

A geotextile is similar to a soil in that it has voids (pores) and particles (filaments and fibers). However, the geometric relationship between filaments and voids is more complex than in soils due to the shape and arrangement of the filaments and the compressibility of the structure with geotextiles. Since pore size can be directly measured, relatively simple relationships between the pore sizes and particle sizes of the soil to be retained can be developed. Three simple filtration concepts are used in the design process: i) if the size of the largest pore in the geotextile filter is smaller than the larger particles of soil, the soil will be retained by the filter. As with graded granular filters, the larger particles of soil will form a filter bridge over the hole, which in turn, filters smaller particles of soil, which then retain the soil and prevent piping (Figure 4) (Holtz *et al.*, 1998), ii) if the smaller openings in the geotextile are sufficiently large enough to allow smaller particles of soil to pass through the filter, then the geotextile will not *blind* or *clog* (Figure 5) (Holtz *et al.*, 1998) and, iii) a large number of openings should be present in the geotextile, so proper flow can be maintained even if some of the openings later become plugged (Holtz *et al.*, 1998).

The following design criteria for geotextiles must be met: i) soil retention, ii) water permeability and, iii) clogging resistance.

To perform effectively, the geotextile must also survive the installation process (survivability criterion) (Zhou, 1998).
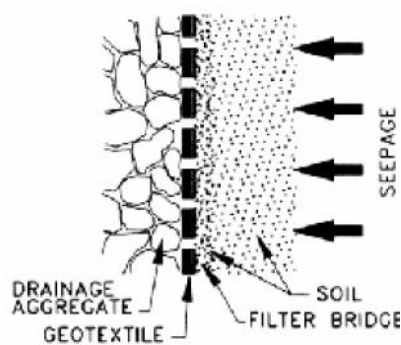

Fig. 4 Filter bridge formation   **Fig. 5** Clogging and blinding.
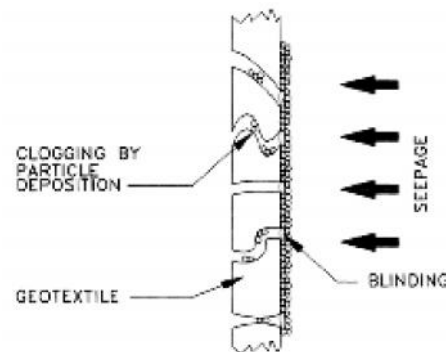
Hydraulic properties are related to the pore size distribution of the geotextile and correspondingly its ability to retain soil particles over the life of the project while allowing water to pass (Lee and Bourdeau, 2006).

The ability of water to pass through a geotextile is determined from its hydraulic conductivity (*coefficient of permeability, k*), as measured in a permeability

test. The permeability of geotextiles can vary immensely, depending upon the construction of the fabric. National and international standards have been set up for the measurement of permeability that is required, most often at right angles to the plane of the textile (crossflow), but also along the plane of the textile (in-plane flow, called *transmissivity*). Albanian Standards specifying those tests are S SH EN ISO 11058 – Water permeability normal to the plane and S SH EN ISO 12958 – Waterflow capacity in the plane (Drejtoria e Përgjithshme e Standardizimit, 2008).

Water flowing is very important for civil engineering earthworks as it prevents the *pressure buildup* of *water*. Coefficient permeability is defined as the product of permeability considering its dimension in the flow direction. The units are nationalized in metres per second. Effectively the coefficient is a velocity, indicating the flow velocity of the water through the textile.

Due to the compressibility of geotextiles, engineers also use a coefficient called the permittivity, which defines the theoretical permeability irrespective of the thickness of the fabric. The *permittivity,*   (permeability divided by thickness), is often determined from the test and used to directly evaluate flow capacity. Permittivity is defined as the cross-plane permeability of a geotextile. In other words, permittivity is a measure of the ease at which water may flow through the geotextile. Permittivity is obtained with S SH EN ISO 11058, by measuring the flow of water, under a given head, moving perpendicularly through a geotextile (Drejtoria e Përgjithshme e Standardizimit. 2008). Values are reported in units of $\sec^{-1}$, allowing for a more traditional value of permeability to be obtained by multiplying the permittivity by the thickness of the geotextile.

Permeability criteria are established based on the critical nature and severity of the application. For noncritical, less severe applications, the permeability of the fabric (k_f) is required to be at least greater than the permeability of the soil (k_s) the fabric is to retain, $k_f > k_s$. For critical-severe applications the fabric permeability must be at least ten times greater than the permeability of the soil, $k_f > 10k_s$ (Karcz and Holtz 1988).

Permeability criteria for woven geotextiles are in terms of the percent open area (POA). When the protected soil has less than 5 percent passing the No. 200 sieve, the POA should be equal to or greater than 10 percent. When the protected soil has more than 5 percent but less than 85 percent passing the No. 200 sieve, the POA should be equal to or greater than 4 percent (Joint Departments of the Army and Air Force, 1995).

The ability of a geotextile to retain soil particles is directly related to *its apparent opening size* (AOS) which is the apparent largest hole in the geotextile. The AOS value is equal to the size of the largest particle that can effectively pass through the geotextile in a dry sieving test. It is also called filtration opening size (FOS).

The procedure for matching a textile to the soil, in order to achieve stability under difficult hydraulic conditions, is to use a textile whose largest holes are equal in diameter to the largest particles of the soil ($O_{90} = D_{90}$). Where hydraulic conditions are less demanding, the diameter of the largest textile holes can be up to five times larger than the largest soil particles ($O_{90} = 5D_{90}$). The relationship between O_{90} and D_{90} is in figure 6 depicted. Particularly difficult hydraulic conditions exist in the soil: i) when under wave attack, ii) where the soil is loosely packed (low bulk density), iii) where the soil is of uniform particle size, or iv) where the hydraulic gradients are high. Lack of these features defines undemanding conditions. Between the two extremes lies a continuum of variation which requires experience and judgment in the specification of the appropriate O_{90} size for any given application (Horrocks and Anand 2000).
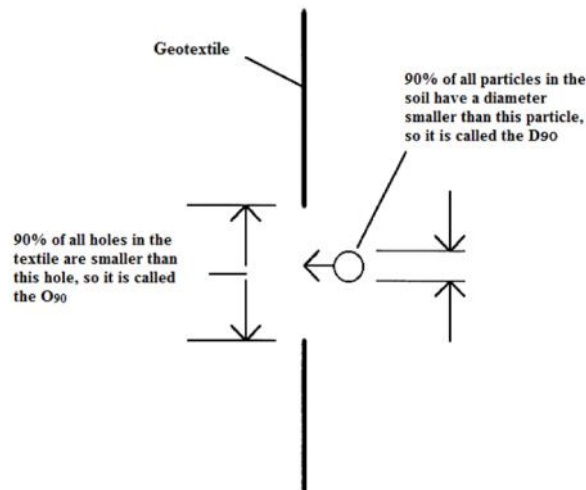

Fig. 6: Relationship between O_{90} and D_{90}

The largest hole sizes and largest particle sizes are assessed by consideration of the largest elements of the fabric and soil. Measuring the largest particles of a soil is achieved by passing the soil through standard sieves. In order to assess a realistic indication of the larger particle diameters, a notional size is adopted of the sieve size through which 90% of the soil passes. This dimension is known as the D_{90} by convention. Similarly, an indication of the largest holes in a textile is taken as the 90% of the biggest holes in the fabric, the O_{90}. Even under ideal conditions, if the O_{90} pore size is bigger than $5D_{90}$, then so called piping will take place. The textile O_{90} pore size should be reduced from $5D_{90}$ towards D_{90} as the ground and hydraulic conditions deteriorate. The current standard in Albania for measuring the characteristic opening size is S SH EN ISO 12956 (Drejtoria e Përgjithshme e Standardizimit. 2008).

4. Situation in Albania

The interest on geotextiles has been lately worldwide increased, especially in developing countries, including Albania. Some applications of geotexiles in Albania are: in the Levan -Dames road (Maccaferri Corporate, 2012), where a non-woven geotextile was placed on top of the pile caps to protect the gravel drain from becoming contaminated by fine soils transported within the ground water. Also in the Gjirokastër –Tepelenë road (Maccaferri Corporate, 2003), where the geotextile is placed at the bottom of the structure in order to save it from the river flow erosion. Moreover, along the cut of the slope was installed a drainage composite to guarantee the drainage of the rainfalls. Both segments are part of the North South Corridor in Albania between Montenegro in the North and Greece in the south. Another application is in the Durrës - Kukës - Prishtinë Motorway, part of the South – East Europe Core Road Network, where the geotextile is placed at the bottom of the embankment to stabilize soils (Gjetvaj *et al.* 2012).

Short and long terms properties of geotextiles are very important for civil engineering and the use of geotextile should be based on standard for every application. Testing geotextiles and evaluating their applicability would be very important for the entire structure where they are applied.

Currently, European standards have been adapted by the Albanian institutions involved in the area. However, their implementation has still to be defined as reviewing the Technical Construction Legislation with regard to some specifics about geotextiles and their use is necessary. Once reviewing has been made, including geotextiles in the list of construction materials would be appropriate. Here, evaluation of conformity is required in a normative way.

Table 1 reports on the standard test methods used to consider the properties here reported as defined by the Albanian Standards based on the European standards.

Table 1. Geotextile properties and their associated S SH EN ISO Standard Test Methods.

Characteristic	Test method	Functions					
[b1]		F*	D*	R*	S*	P*	B and STR*
Tensile strength	S SH EN ISO 10319	X	X	X	X	X	X
Elongation at maximum load	S SH EN ISO 10319		X	X		X	X
Static puncture resistance (CBR test)	S SH EN ISO 12236			X	X	X	X
Dynamic perforation resistance	S SH EN ISO 13433	X	X	X		X	X
Water permeability normal to the plane	S SH EN ISO 11058	X					
Characteristic opening size	S SH EN ISO 12956	X					
Water flow capacity in the plane	S SH EN ISO 12958		X				

* F – filtration; D – drainage; R – reinforcement; S – separation; P – protection, B and STR – fluid barrier and strength release.

5. CONCLUSIONS

Geotextiles properties make them fit for civil engineering use. Mechanical and hydraulic properties are the most important ones.

Mechanical properties refer to the ability of a textile to perform work in a stressed environment and to resist creep failure under any given load condition. They provide appreciative geotextile strength and/or compressibility under varying loads. The most important are tensile strength and tensile creep. Load-strain and the complete stress-strain curve are closely related. Creep can cause the physical failure of a geotextile if it is held under too high a mechanical stress.

Hydraulic properties relate to the ability of geotextile to retain soil particles over the life of the project while allowing water to pass. The ability of water to pass through a geotextile is related to characteristics such as permeability, permittivity, transmissivity, while retaining soil particles is related to the Apparent Opening Size of geotextiles and its relation to the particle soil size.

The data here reported could be used to develop a properties classifying system in order to define geotextiles for routine applications in Albania. The classification system was carried out for the first time by the French Committee of Geotextiles and for the design of a number of routine geotextile applications. This system has been used by other countries to develop their own systems.

REFERENCES:

Christopher BR, Holtz RD. 1985. *Geotextile Engineering Manual.* FHWA-TS-86/203, Federal Highway Administration, Washington, D.C., USA.

Drejtoria e Përgjithshme e Standardizimit. 2008. *Katalogu i Standardeve Shqiptare*. Tiranë, Shqipëri.

Gjetvaj V, Bušeli L, Šimovi J, Ulker S. 2012. *Geotechnical problems on the Section 3 on the Motorway in Kosovo.* The first Albanian Congress on Roads, Tiranë, Albania.

Holtz RD, Christopher BR, Berg RR. 1998. *Geosynthetic Design & Construction*

Guidelines. FHWA Technical Report No. FHWA-HI-95-038, Federal Highway Administration, Washington, D.C., USA.

Horrocks AR, Anand SC. 2000. *Handbook of Technical Textiles.* Woodhead Publishing Limited in association with The Textile Institute. Cambridge, England.

Joint Departments of the Army and Air Force. 1995. E*ngineering Use of Geotextile*. TM 5-818-8/AFJMAN 32-1030, Washington, D.C., USA.

Karcz DA, Holtz RD. 1988. *Joint Highway Research Project Jhrp-88/14. Development of the Idoh Classification system for Geotextiles.* Purdue University, West Lafayette, Indiana, USA.

Maccaferri Corporate. 2003. *Environmental/ hydraulic & erosion control/ slope protection. Gyrocaster-Tepelene Road, Albania,* Officine Maccaferri S.p.A., Via Agresti, 6 - P.O. BOX 396 – 40123, Bologna, Italy.

Maccaferri Corporate. 2012. *Major Highway embankment reinforcement. Levan to Dames Road, Albania.* Ref: INT/CH/ALB/BP001*,* Officine Maccaferri S.p.A., Via Agresti, 6 - P.O. BOX 396 – 40123, Bologna, Italy.

Sangho L, Bourdeau PL. 2006. *Filter Performance and Design for Highway Drains.* FHWA/IN/JTRP-2005/1, Purdue University. West Lafayette, Indiana, USA.

Zhou Y. 1998. *Geosynthetic Filters*. Continuing Education and Development, Inc. NY 10980, USA.

Zornberg JG, Thompson N. 2012. *Application Guide and Specifications for Geotextiles in Roadway Applications*. Center for Transportation Research, The University of Texas at Austin 1616 Guadalupe, Suite 4.202, Austin, TX 78701, USA.

HEALTH THROUGH NUTRITION: FACTORS OF NUTRITION AND THEIR IMPACT ON MORBIDITY

November 21, 2015

Making Science work for the benefit of the Society

The Section of Natural and Technical Sciences of the Albanian Academy of Sciences (hereafter reffered to as Academy) organized on November 21st in close partnership with the University of Medicine of Tirana, Ministry of Health, Institute of Public Health and Compulsory Health Insurance Fund, the Scientific Conference "Health through nutrition: factors of nutrition and their impact on morbidity". As the Academy through its Section aims at introducing new initiatives in the area of research and education to suppor t current needs of the country and helping find solution to many crucial issues concerning the development of the country by providing high state institutions with appropriate advice and expertise, the conference run in line with the Government Program and aimed at: i) evidencing the nutrition's impact on morbidity of the population in Albania and prevention from specific diseases, ii) involving all the stakeholders and raising awareness on the role the food palys in morbidity reduction, iii) exchaning inforamtion, buiding and raising the capacities among the stakeholders involved in the area.

The beneficary target groups were the Ministry of Health of Albania, Compulsory Health Insurance Fund, Regional Health Directories, Primary Service of Family Doctors, wide population and food business.

At the end of the Conference, the Conference Declaration Chart was approved and recommandations to state institutions involved in the area were made.

ALBERT EINSTEIN AND THE GENERAL THEORY OF RELATIVITY-IN THE FRAMEWORK OF SCIENCE EDUCATION PROGRAM

November 27, 2015

2015 is called the year of the festival of light because it marks the 100th anniversary of the General Theory of Relativity by Albert Einstein, one of the most eminent personalities of all times. Consequently, the Section of Natural and Technical Sciences organized in the framework Science Education Program, the International Conference "Albert Einstein and the General Theory of Relativity".

He is the fuzzy haired genius who unlocked the secrets of the universe. He was also a great peacemaker. Thanks to Albert's great $E = mc^2$ equation, the world has nuclear power, radiation therapy, and gamma ray scans. His explanation of the Photoelectric Effect has led to fluorescent lights, automatic doors, lasers, TV, computers and much more. Children, also have to be grateful to him any time they play video games.

A list of guests included Mr. Dhori Kule, Rector of the University of Tirana, Albania, Ligor Nikolla, Dean of the Faculty of Ingeneerin mathmatics and Physics, Mynyr Koni, Dean of the Faculty of Natural and Technical Sciences, Academitians, professors of physics and students, members of the Albanian parliament and many representatives of higher education and research institutions and His Excellency Boaz Rodkin, Ambassador Extraordinary and Plenipotentiary of Israel in Albania.

Of great interest were the papers of Prof. Halil Sykja, Prof. Dritan Spahiu, Prof. Dr. Mimoza Hafizi, Dr. Lorenzo Amati from the Italian Institute for Astrophysics who shed light on some of the achievements in the area of physics and astrophysics due to Albert Enstein's theory.

In the end, Dr Peter McGrath - Coordinator of the *Science Diplomacy* Program run by the World Academy of Sciences (TWAS) for the advancement of science in developing countries provided information about the legacy of Einstein as diplomat scientist.

IN VITRO CULTURE – FROM SCIENTIFIC RESEARCH TO MARKET; THE NEED FOR A VIRTUAL NATIONAL NETWORK OF PLANT BIOTECHNOLOGY

December 3, 2015
Hall "Aleks Buda", Albanian Academy of Sciences

The Academy of Sciences is dedicated to promoting the exellence in science, its benefits to the society and providing scientific advice for policy. Consequently, the Section of Natural and Technical Sciences of the Albanian Academy of Sciences organised in close partnership with the Ministry of Agriculture, Rural Development and Water Administration; the Center of Agricultural Technology Transfer, Vlora; University of Tirana and Agricultural University of Tirana; Academy of Sciences and Arts of Kosovo and the University, Ukshin Hoti, Prizren, Kosovo the Scientific Symposium "*In vitro* culture – from scientific research to market; the needs for a Virtual National Network of Plant Biotechnology". The Symposium aimed at araising awareness on the relationship between technlogic development and products and services of high quality in line with Government Program.

The main topics adressed were as following: i) phytosanitary status of fruit trees in Albania, production of virus-free saplings applying the meristem culture technique and making certified material available to farmers, ii) current diagonstic opportunities for phytopatologies of cultivated plants applaying molecular methods, iii) application of tissue culture as an alternative method of medium and long term conservation of plant germplasm and a variability source for genetic improvement and breeding of cultivated plants, iv) use of *in vitro* technique as a means to address genetic plant improvement to create new plant forms of high quality and, v) improving and fostering biotechnological research in the area of *in vitro* culture in Albania by creating the Virtual National Network of Plant Biotechnology; financial support intented to be over-arching, in order to support the specific scientific areas would be of crucial importance.

In the area of biotechnology, Albania is characterized by an excellent infrastructure and well-equipped research facilities with highly qualified staff: i) scientific laboratories of cell and tissue culture and molecular biotechnolog*y*, Faculty of Natural Sciences, University of Tirana, set up by budgetary funds. Up to 2007, these laboratories were assets of the Institute of Biological

Research, Albanian Academy of Sciences. Since then, these laboratories have been infrastructure for Bachelor, Masters and PhD programs and thesis in the area of *in vitro* technology of germplasm of spontaneous species which are an interesting component of Albanian flora (endemic, rare, threatened and with high economic interest species as the medicinal-aromatic plants, authocthonous fruit trees species etc.), identification of plants viruses, evalutation of genetic polymorphism in subspecies and interspecific level, and the evidence of somaklonal variations of *in vitro* plants. The skillfull staff make this Department very important for the University of Tirana, ii) the complex of scientific and commercial laboratories with greenhouses for production of fruit seedlings (former Institute of Pomology, Vlora, currently the Center of Agricultural Technology Transfer), which primary purpose is the certification of fruit species rootstocks (in the framework of Italy-Albania Project in the collaboration with CIHEAM (Mediterranean Agronomic Institute), Bari, Italy, in the framework program INTERREG III, 2001-2011, of the European Union (EU). These laboratories were set up with substantial funds to shore up the support for the faculty. Up to 20 thousand certified seedlings of fruit species rootstocks were produced up to 2011, when the production stopped, iii) laboratory of plant genetic bank, State Agency for Seeds and Saplings (Project of the World Bank in the collaboration with the Project of Agricultural Services, ASP) (2003 – 2007), which is part of the Plant Genetic Resources Center since 2007. Unfortuantely, the laboratory has not been put into function, yet, regardless considerable investments and the efforts made to put the laboratory into function., iv) scientific and didactic laboratory, Department of Agricultural Production, Faculty of Agriculture and Environment, Agriculture University of Tirana (TEMPUS Program) (2011-). Eventhough the laboratory is well-equipped, further training would be necessary for a better scientific research in the field of *in vitro* cultures and, v) didactic laboratory, Department of Horticulture and Landscape Architecture, Faculty of Agriculture and Environment, Agriculture University of Tirana (project of Ministry of Education and Sports, UNILAB), which additional equipment and specialized staff are of fundamental importance.

In his welcome speech, the Head of the Section of Natural and Technical Sciences, Acad. Prof. Dr Salvatore Bushati emphasized the impact of technologic development on biotechnology and Science education Program. In addition to prof. Bushati, Prof. Dr. Ramë Vataj from the University, Ukshin Hoti, Prizren, Kosovo; Prof. Dr. Mynyr Koni Dean of the Faculty of Natural Sciences, University of Tirana; Prof. Dr. Ardian Maçi by the Faculty of Agriculture and Environment, Agricultural University of Tirana and Mr. Flamur Hysi on behalf of the Ministry of Agriculture, Rural Development and Water Administration welcomed the participants to the conference. In this Session,

Acad. Ass. Efigjeni Kongjika held main presentation "*State and perspective of scientific research in the field of plant biotechnology (in vitro culture) in the World and Albania*".

After presenting the orchard phytosanitary situation and the risk of viral diseases spread in nursery in Europe and Albania by various specialists from several institutions (CERTIS Company, Italy; CIHEAM, Mediterranean Agronomic Institute, Bari, Italy; Faculty of Agriculture and Environment, Tirana Agricultural University and Faculty of Natural Sciences, Tirana University), the specialists of the Center of Agricultural Technology Transfer, Vlora etc. presented scientific data on the *in vitro* culture application for the production and micropropagation of sanitated plant material. Aspecial attention was paid to the presentation of these specialists with regard to the possibilities of introducing sanitated biotechnologic product in the Albanian market and to constructe new private commercial laboratories.

Many specialists of the Academy of Sciences, University of Tirana, Agricultural University of Tirana and the University, Ukshin Hoti, Prizren reported their long experience in the field of *in vitro* cultures on the micropropagation, meristem culture and *in vitro* conservation of plant germplasm. The possibilities of using new *in vitro* techniques and the perspective of use of modern biotechnologies at the laboratories of molecular biotechnology and *in vitro* cultures of the University of Tirana and Agricultural University of Tirana.

A group of specialists from different institutions emphasized the need for a National Virtual Network of Plant Biotechnology as a means to address contemporary research in the field of plant biotechnology. The network would play the role of an *ad-hoc* committee where some of the most eminent personalities would come together to discuss the long-term needs, and tendencies of biotechnology.

At the end of the symposium, a chart with recommendations from the auditor to the scientific community and policymakers was approved.

ANDREA PIERONI AWARDED HONORARY MEMBER OF THE ALBANIAN ACADEMY OF SCIENCES

December 7, 2015

The Section of Natural and Technical Sciences of the Albanian Academy of Sciences is honored and priviledged to announce Prof. Dr Andrea Pieroni Honorary Member of the Academy of Sciences of Albania.

Prof. Andrea Pieroni, after receiving his masters in Pharmacy and trained in Medical Botany at the University of Pisa, earned his doctorate in Life Science, University of Bonn, Germany. A Research Assistant at the University of London since 2000, as well as Lecturer and Senior Lecturer at the School of Life Sciences of the University of Bradford in Northern England since 2003, he returned to Italy in January 2009.

Currently, he is a Professor for the Environmental and Applied Botany subject area (food botany, ethnobotany, and ethnobiology), University of Gastronomic Sciences, Pollenzo/Bra, Italy. Professor Pieroni has been the President of the International Society of Ethnobiology and currently he is Editor-in-Chief of the *Journal of Ethnobiology and Ethnomedicine*. He is also a board member of several International Ethnoscientific Societies and publication editors, as: *Journal of Ethnobiology and Ethnomedicine* (editor-in-chief), *Journal of Ethnopharmacology, Journal of Ethnobiology, Ethnobotany Research and Applications, Food and Foodways, Ethnobiology and Conservation*.

He has been the P.I. of the first collaborative research project funded by the EU Commission focused on traditional plant knowledge in the Mediterranean (RUBIA, 2003-2006). His research focus is on gastronomic and medical ethnobotany in the Mediterranean area, the Balkans, and among migrant communities in Western Europe.

His courses are focused on: i) Food Biodiversity Sciences (in English), ii) Ethnobotany and Ethnobiology (in English), iii) Plant Sciences (in the Italian language), iv) Molecular basis and physiology of taste and Nutraceuticals (in the Italian language) and, v) Ecology(in the Italian language).

His main research topics are: i) ethnobiology: cross-cultural food and medical ethnobotany, ii) Ethnomedicine: folk medical practices, iii) human ecology: local

ecological knowledge and its mechanisms of adaptation and, iv) ethnoveterinary among pastoralists.

Studies areas/communities concern linguistic and/or religious minority groups and cultural boundaries in South and South-Eastern Europe and migrant communities in Western Europe

Research areas concern ethnobotany, ethnobiology, ethnomedicine, ethnoveterinary, social anthropology, migration studies, ethno-ecology, human ecology, gastronomy, ethnozoology, Balkan studies, Albanian studies, folk medicine, local gastronomy, medicinal anthropology, food anthropology, linguistic minority, Rumania (anthropology), gastronomical literature, gender, culture, emmigration and diaspora, ethnobiological studies, ethnobiological, ethnobiological for Albania.

Community Studies involve linguistics and religious minority groups, and their cultural border in South/South-East Europe and immigrant communities in Western Europe, as well.

His research focus is on Ethnobotanic Medicine and in the Mediterranean, Balkan and Immigrant communities in Western Europe. Current studies center on plant use as medicine and food, from the linguistic minorities in Southern Italy, Balkan and in immigrant communities (Eastern Europenians, South Asian, Turkish etc.) in Europe. He was a leader and pioneer in the first research project financed from the European Commission, focused on traditional medicinal knowledge in the Mediterranean (RUBIA- Ethnobotanical and ethnographic inheritance of the Mediterranean traditional technology, tools and uses of neglected and wild plants for food, medicine, textile and coloring) 2003-2006.

PROF. DR NIKO PANO AWARDED
"HONOR OF THE ACADEMY"

December 22, 2015

The Albanian Academy of Sciences based on the proposal from the Section of Natural and Technical Sciences awarded Prof. Dr Niko Pano "Honor of the Academy".

Niko Pano is an eminent personality in the area of hydrology, internationally recognized as well. His early work in hydrology and the *sub-fields* is related to water and energy economics. In addition, he is well known for his efforts to foster intellectual pursuits of the staff of the first Institute of Hydrometeorology in Albania in the area of scientific research.

He is the author of the "Water resources of Albania", published by the Albanian Academy of Sciences.