International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-5, Issue-1, May 2016

Statistical Landslide Susceptibility Mapping by Using GIS-based Weight-of-Evidence (WofE) Analysis in Takhar Province of Afghanistan

Mohammad Kazem Naseri, Dongshik Kang

Abstract— **Takhar province in Afghanistan is a susceptible region for the occurrence of landslides, debris flow, earthflow, rock falls, soil creeps, and loess failures. Recent huge mass movement in Abi-Barik and Rushtaq villages of Badakhshan and Takhar provinces respectively, have caused numerous loss of lives and damaged properties. The problem occurred due to the direct influence of terrain types and their characteristics. Moreover, seismicity, tectonic faults, and rainfall have all contributed to landslide occurrences. In our study area, we created a landslide susceptibility map by means of a statistical analysis using the weight of evidence (WofE) method to designate and identify the landslide susceptibility zone (LSZ). The LSZ is at a scale of 1:200,000 and present the distinct susceptibility classes. For this purpose, we used 30m resolution digital elevation model (DEM) data and GIS software to prepare causative parameters that mostly contribute to landslide initiation and glacier mass movement. These parameters are namely, slope angle, slope aspect, geology, landcover type, soil depth, precipitation, topography, stream distribution, and proximity to the earthquake epicenters and tectonic faults. The final LSZ map was classified into five zones of landslide susceptibility levels, such as: very low (8%), low (17%), medium (19.5%), high (33%) and very high (22.5%) of the total area. Based on the landslide locations and input parameters the final result has been validated by the receiver operating characteristic curve (ROC) analysis. According to the plot, the model shows 85.4 percent prediction accuracy (area under the curve=0.854). This is not a proven perfect prediction for specific site improvement but may serve as a suitable model for regional scale development.**

Index Terms— **Takhar Province of Afghanistan, Statistical Landslide Susceptibility Mapping, GIS and Weight of Evidence.**

I. INTRODUCTION

The study of landslides has drawn worldwide attention mainly due to increasing awareness of its socioeconomic impact as well as the increasing pressure of urbanization on the mountain environment (Aleotti and Chowdhury 1999) [1]. Landslides represent a serious threat to human life and their socio-economical activities in most of the high mountain chains. The recognition of landslide prone topography is becoming increasingly important in environmental management and land use decisions and is vital for regional as well as zonal and local community planning. (H. Vijith et al, 2013) [2]. Among the recent approaches for landslide hazard mapping, the statistical approach is considered to be more suitable for landslide hazard mapping over large and complex areas (Xiwei Xu* et al, 2012) [3]. Statistical classification methods have developed rapidly in recent years and become an effective tool to determine landslide susceptibility over large and complex areas (Geological Survey of Ireland, 2011) [4]. Such methods provide quantitative estimates of “where” landslides are expected, based on detailed information on the distribution of past landslides and a set of thematic environmental information. Earthquake-triggered landslides always caused tragic deadly events and serious economic losses (Keefer, 1984) [5].

Afghanistan is a landlocked mountainous country located within South and Central Asia. The country is prone to natural hazards such as flood, earthquake and landslides. North-eastern of Afghanistan is one of the most susceptible regions to loess failure, landslide, and rock fall. Because of the slope gradient, complex seismic activity, and other causative geological factors, recent huge mass movement in Abi-Barik and Rushtaq villages of Badakhshan and Takhar provinces respectively, have caused enormous loss of life and property damages. The fatalities were reported to be more than 300 houses destroyed and about 400 people killed. But some local reporters said almost 1000 persons have died (News and Local Official) [6].

In the northeastern of Afghanistan, as well as the larger rockslides and rock slide complexes there are categories of fine-grained mass movement that occur where either water or seismic acceleration are the primary causes of slope failure(John F. Shroder et al, 2011)[7]. Northeast of Afghanistan is an exceptionally high seismic region (Russell L. Wheeler and Kenneth S. Rukstales) [8]. Historically, 21 earthquake epicenters with magnitudes of 7 or more have been located in or near this area. In the case of such adversity following an earthquake were landslide happens, a landslide susceptibility analysis would be essential in classifying areas of low to high hazard levels to save lives. However, there are still too many problems in term of data preparation, and landslide characteristic analysis in the area. Landslide susceptibility mapping is diverse, numerous and often vary dependent on the country, landscape, purpose and finances available for the project.

The objective of this study is to create a landslide hazard zonation map scale at 1:200,000 for the proposed study area. Consequently, the first step was collating a landslide inventory map simply using Google Earth, 10m contour lines and GIS software. The second step, all the ten input causative factors such as slope gradient, soil, slope aspect, geology, landcover, stream distribution, fault location and distance, precipitation, seismicity and relative relief have been prepared in the same areal extent and pixel size by using 30m resolution DEM. The geology and tectonic maps of Afghanistan were also used. These data mostly contribute to the information on landslide initiation, debris flow, earthflow,

Mohammad kazem naseri, Graduate School of Engineering and Science, University of the Ryukyus, Okinawa, Japan, 05037166417.

Dongshik Kang, Information Engineering, University of the Ryukyus, Okinawa, Japan, 09019499546.

rock fall, avalanches, and soil creeps.

II. STUDY AREA

The study area has covered central and southern part of Takhar province which is 9,543 square kilometers. Takhar province is located in the North part of Afghanistan within the latitude of 36.4011°N, longitude of 69.2842°E. It is a mountainous region with elevation ranging from 486m to 5,813m. Slope within the area is ranges between 0 to 81 degrees. In the study area, a total of 247 landslides were mapped which were classified as to types and characteristics, such as soil creep, debris avalanches, earthflow, fluvial undercutting, glacier activities, rock fall, rotational and translational landslides. The original landslide dataset has been divided into two parts: one part to be used to model the spatial structure and produce a surface raster map, the other to be used to compare and validate the output surface raster map. As a result, 138 landslide locations were used to produce the surface raster map in the model and 109 landslide locations were used to compare and check the model accuracy (Fig. 2).

Takhar is vulnerable to an earthquake-induced landslide. Most of the time, even the low magnitude earthquake can cause serious property damage and loss of lives. Some of the previous events were the earthquake of 4 February 1998 that occurred in Rustaq district, in the remote Takhar province of north-easthern Afghanistan, at 19.03 local time, which according to the United States Geological Survey (USGS), measured 6.1 on the Richter scale. The depth of the earthquake was 33 kilometers. The earthquake has levelled 28 mountain villages and killing at least 4,500 people and Thousands more were deemed missing or buried because the village houses were typically made with mud walls, flat timber, and mud roofs. These structures collapsed as a result of the shock and the following landslides (Report of the UNDAC Mission 10 February to 6 March, 1998) [9]. On 13 April 2014, an unusual landslide occurred at Rustaq district of Takhar province and resulted in four people dead and dozens of houses destroyed. According to the United States Geological Survey, a minor 4.1 magnitude earthquake struck northeast of Afghanistan at 2am at a depth of 203 kilometers. A huge landslide happened at the location of 36.514°N, and 70.403°E that was caused when the earthquake struck (Local news and media).

Geology is complex within the study area as it hosts Intrusive rocks (Granite, Granodiorite, Granosyenite, Granophyre, Andesite, Diorite, Plagiogranite, Dunite, Peridotite, and Serpentinite) and stratified formations of geological ages(Miocene, Jurassic, Cretaceous, Pliocene, Proterozoic, Quaternary, Eocene, Karachatur horizon, Tournaisian sub- stage, Maestrichtian-Paleocene, Namurian stage, Ordovician, Rhaetian, Silurian-Devonian) with material types of (Red and brown clay, siltstone, sandstone, conglomerate, limestone, clay, gypsum and coal, red sandstone, acid volcanic rocks, Variegated sandstone, biotite and garnet-biotite gneiss and schist, quartzite, marble, amphibolite, mudstone, shale, shingly and detrital sediments, gravel, sand, loam, loess, travertine, andesitic basalt, and olivine basalt).

Landcover types of the study area is diverse from natural vegetation to Intensively cultivated agricultural land such as: rain fed crops (flat lying areas), rangeland (grassland, forbs and low shrubs), rain fed crops (sloping areas), marshland permanently inundated, intermittently cultivated, fruit trees, pistachio forest, natural forest (open cover), degenerate forest/high shrubs, vineyards, gardens, permanent snow, water bodies, rock outcrop, bare soil and settlements.

Fig 1 is an aerial view of a landslide caused by the massive earthquake, in Rustaq district of Takhar province, Afghanistan, 12 April 2014 at the location of 36.514°N and 70.403°E. At least four people have been found dead and more than 100 houses were damaged when an earthquake jolted two villages early 12 April in Takhar. The photo(Fig. 1) has been taken by: EPA/NASIR WAQIF.

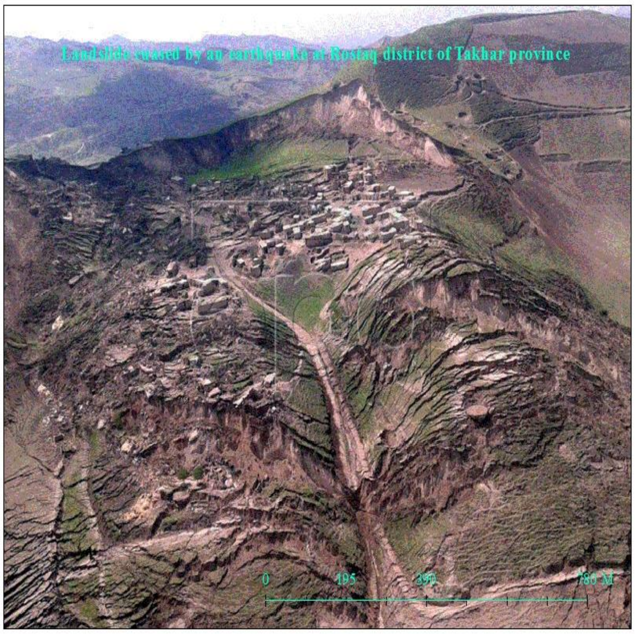

Fig. 1: Landslide in Takhar 12 April 2014

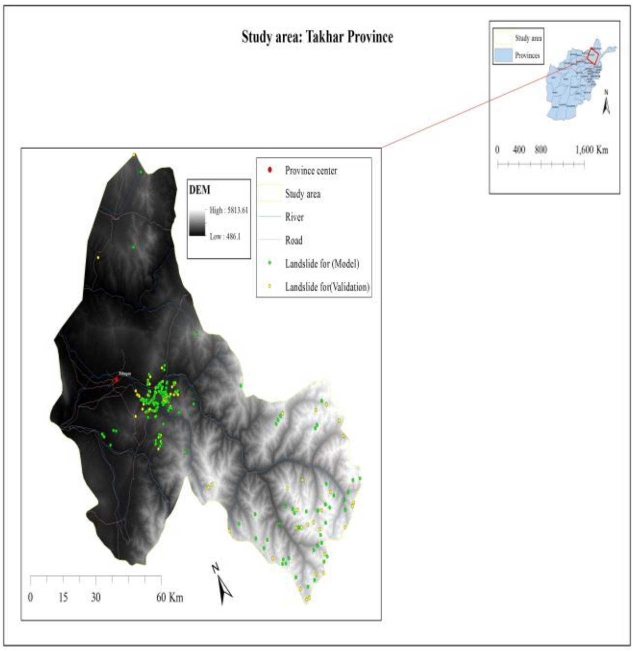

Fig. 2: Study area map

III. METHODOLOGY

Landslide inventory mapping is the first stage in any landslide susceptibility mapping and study (Fausto Guzzetti, et al, 2012) [10]. The correct identification and classification are fundamental in ensuring the accuracy and usability of a susceptibility map.

In the methodology, we used the WofE formula which is based on pixel values of each raster to calculate the information values. The ten independent causative factors such as landcover, precipitation, slope, aspect, soil, geology, elevation, seismic risk, proximity to the tectonic faults and streams were used in the model. The importance of the statistical method is that we can consider all of the contributors which thought to be causative for the landslide occurrence as weighted classes on the model. There are many landslide mapping methods, but recently statistical method is becoming common and an interdisciplinary approach among the researchers because of its suitability for a regional study and analysis of landslide initiation. The method was developed by Yin and Yan (1998) which since that time it has been used in mineral exploration as well as landslide mapping. The calculation applied to this hazard information method is based on the following formula, as for the calculation the information value I_i for variable X_i:

$$I_j = \log \frac{S_i / N_i}{S / N} \qquad (1)$$

Where:

S_i the number of landslide pixels in each class and the presence of variable X_i,

N_i the number of landslide pixels in the map with variable X_i,

S the total number of pixels in each class,

N the total number of pixel in the map.

The degree of hazard for a pixel *j* is calculated by the total information value I_j:

$$I_j = \sum_{i=1}^{m} X_{ij}.I_j \qquad (2)$$

Where, *m* is the number of variables

X_{ij} is *0* if the variable X_i is not present in the pixel *j* and *1* if the variable is present

The bigger the I_j value is, the more unstable pixel *j* within the area.

IV. THE CAUSATIVE FACTOR SELECTION

Many causative parameters could contribute to the landslide initiation. As for the study area the most contributor parameters are: seismicity of the area, proximity to the tectonic faults and streams, slope gradient and aspect, precipitation, geology, landcover and soil types, and releif classes. Most of the input data layers were derived and have been prepared using 30-meter resolution DEM, but for the geology, earthquake epicenters, soil and landcover, USGS survey and documented files, Ministries of Agriculture, and Mines and Petroleum data have been used.

Seismicity and Tectonic Faults: seismicity within the area is growing rapidly. Regarding tectonic processes, the estimated peak ground acceleration for Takhar with an exceedance probability of 10% in 50 years is between 5 to 8 ms-2. Therefore, distances from the historical earthquake that have happened in the past, buffered in GIS as (0.5, 2.5, 5, 8 and >8) kilometer distance classes. Tectonic faults have been considered by generating buffer distances (0.5, 3, 6, 10, 15, 25, and >25) kilometers from the faults locations.

Topography and Precipitation: Topography is one of the important factors for landslide occuring, especially, for debris avalanches and rocks falling. It is sliced from DEM in (500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, and >4000 meters elevation) classes. Based on the USGS global weather data station, the precipitation ranges from low of 500mm to a high of 991mm, annually. We reclassified the precipitation into classes at 500, 600, 700, 800, 900, and 991 mm for the modeling and analysis.

Soil depth: the soil varies from shallow, moderately deep to deep and very deep within the study area. Mostly, the soil layer is located as a shallow cover at the top of the bedrock layer. Therefore, in cases of high precipitation and the slipperiness of the bedrock layer aided by the gravitational force enhances landslides occurrences.

Slope and Aspect: Slope degrees were derived from DEM data and reclassified using GIS into classes at 10, 20, 30, 40, and >40 degrees. The various slope face orientation (i.e. Flat, North, Northeast, East, Southwest, West, and Northwest) was used in the modeling. Due to high glacial events and glacial drifts, landslide or debris avalanches formed some lakes. These lakes during a period of time and precipitation could facilitate a huge mass movement down the slope. Therefore, in order to consider the witness index of the study area, we generated buffer distances from the streams at 0.5, 1.5, 3.5, 6, 9, and >9km classes.

Landcover type: The area covered were mostly classified as: rock outcrop bare soil, intensively cultivated (2 crops per year), rain fed crops (flat lying areas), rangeland (grassland, forbs, and low shrubs), rain fed crops (sloping areas), intensively cultivated (1 crop per year), marshland permanently inundated, fruit trees, intermittently cultivated, settlements, permanent snow, gardens, pistachio forest, vineyards, natural forest (open cover), degraded forest, high shrubs, and water bodies.

Geology: Geology of the study area has been classified according to geological ages such as: Quaternary (early, middle, late and recent), Pliocene and Miocene (early and late), Proterozoic, Cretaceous Undifferentiated (early, and late), Jurassic, Triassic(early and late), Mississippian, Namurian stage, Lower and Upper Tournaisian sub stage, Eocene, Paleocene, Karachatur horizon, Pennsylvanian, Silurian-Devonian and Ordovician.

All the causative parameters have been classified into different classes and converted to the same raster pixel resolution and aerial extent with the same coordinate system.

V. STATISTICAL LANDSLIDE SUSCEPTIBILITY

Landslide is a natural global phenomenon that no society can prevent it from the happening, unless the reasons for landslide initiation and the factors may cause those reasons are investigated and scientific work has been done to recognize and mitigate the risks. Takhar is located within a very active seismic zone where there have been cases of landslide events that were triggered by an earthquake in recent years. For this reason, site selection and identification of landslide prone areas; use of the landslide susceptibility mapping and hazard zonation were the primary approaches. However, for specific site selection, there is the need for a more detailed investigation based on the infrastructure type which is going to be built in the area, but for a regional investigation and planning, a susceptibility mapping plays a vital role.

In landslide susceptibility evaluation, limited data has always influenced the accuracy of the analysis. Landslide Inventory map is an essential tool for the landslide susceptibility mapping. Unfortunately, the landslide inventory map was not available for this study area. Hence, for this study, a landslide inventory map was created using Google Earth images and supported by 10-meter intervals contour lines.

In this study WofE method has been used which of the total landslide pixels were calculated on each class of the causative parameters and then the overall area probability has been considered individually for every input parameters. Overall, ten independent causative factors have been cross referenced with landslide inventory map in order to develop the final LSZ map at scale of 1:200,000(Fig 3). The final LSZ map has been classified into five zones of susceptibility levels, which were: very low (8%), low (17%), medium (19.5%), high (33%) and very high (22.5%) of the area. The Southeast part of the study area showed to be a very high susceptible region to landslide initiation, which may be due to glacial activities. Based on the landslide locations and input parameters the model performance result was validated and showed 85.4% prediction accuracy. The ten evidential theme have been used to create the LSZ map and each parameter showed positive correlation for landslide occurrence. But, classes such as permanent snow, approximately 1,000 mm precipitation annually, the geological formation of early Triassic, higher elevation, and stream distances have the higher positive weights (Table 1).

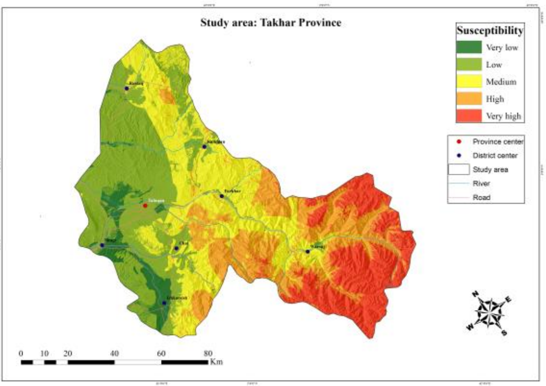

Fig. 3: Susceptibility classes

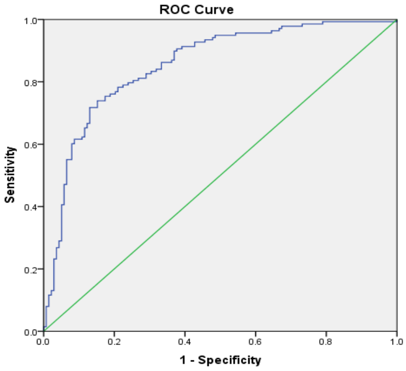

Fig. 4: AUC constructed by ROC method

Table 1: Classified factors with positive weight

Factors	**Classification**	W^{+}
Stream distance	3.5km	0.11
	6km	0.08
Slope degree	20	0.30
	30	0.07
Fault distance	0.5km	0.30
	1km	0.20
Seismicity	2.5km	0.14
Precipitation	991mm/year	0.69
Soil	shallow	0.41
Relative height	500m	0.32
	4000m	0.34
	5813m	0.64
Aspect	North	0.14
	Northwest	0.03
Landcover	Rock Outcrop and Bare Soil	0.35
	Permanent Snow	0.84
Geology	Late Cretaceous	0.16
	Late Jurassic	0.44
	Early & Middle Jurassic	0.15
	Early Triassic	0.66
	Early Proterozoic	0.52

VI. VALIDATION

Model accuracy and its performance depend on the causative parameters used to generate the final LSZ map. In this study, receiver operator characteristic (ROC) curve analysis used to validate the model performance. It is available in the IBM SPSS Statistics 20 software, first released in the USA in 1968. The ROC curve is a graphical plot of false positive rate (1-specificity) on the X-axis and true positive rate (sensitivity) on the Y-axis at different cut-off values. ROC graphs have long been used in signal detection theory to depict the trade off between hit rates and false alarm rates of classifiers (Egan, 1975; Swets et al., 2000). ROC analysis has been extended for use in visualizing and analyzing the behavior of diagnostic systems (Swets, 1988). The accuracy of a test is measured by the area under the ROC curve (Williams et al. 1999) [11] which is known as AUC. Theoretically, the AUC ranges from 0.5 up to one. The value of 0.5 is known as the reference line which means as a worthless prediction but as much as the value is getting closer to the top left corner of the plot near to value of 1 represent the higher prediction accuracy of the model. In the present study, the ROC curve has been constructed by using the 109 landslide locations that have not been used in creating the model itself but to validate the model. The plot shows the AUC of 0.854 which is not a perfect test result but still the model performance is acceptable as a good indicator showing 85.4 percent prediction accuracy (Fig.4).

VII. DISCUSSION AND CONCLUSION

Landslides are the common phenomenon in Takhar province of Afghanistan. They vary from soil creeps, rock debris flow and avalanches, earthflow, glacial deformation, loess failure, rotational and translational landslides. Recent landslide incidence has caused enormous damages to the life, property and infrastructure in Takhar and whole northeast of Afghanistan. Therefore, landslide prone areas should be identified in advance in order to mitigate such damages. An LSZ map scale of 1:200,000 is the primary objective of this work covering the study area in central and southern part of Takhar province. Using WofE method in combination with a GIS application for regional statistical landslide susceptibility mapping have rapid grown in recent years. GIS can play a vital role in data acquisition, vulnerable area identification, landslide risk reduction, disaster management and recovery. The results' accuracy directly depends on the causative parameters analysis and data preparation in GIS. In the present study, WofE method and expert-based knowledge permit for area identification and LSZ map generation. The ten independent terrain variables were used as input to the model and special relationship between evidential themes and landslide locations were calculated statistically. Among the causative factor classes, areas with shallow soil which is facing north and northwest within slope angle 20-30 degrees, approximately located up to 3 kilometers to the tectonic faults and earthquake epicenters with over 900 mm precipitation annually, are the most susceptible terrains for landslide initiation. Geologically, the formation of early, middle and late Jurassic, undifferentiated late cretaceous, and early Triassic and Proterozoic, especially with permanent snow landcover are the most contributors for landslide initiation. The 85.4% accuracy obtained from the validation using ROC curve and indicate the best model performance and prediction within the study area. The result of this study could not reflect the best prediction and assumption in terms of the specific site and area development. Due to the small scale (1:200,000). However, the study results can serve as a primary information for site selection, infrastructures development and disaster management on a regional scale. The engineers, developers, and designers should consider the suitability of small scale topographic conditioning for the landslide assurances in advance.

ACKNOWLEDGMENT

We gratefully acknowledge the Graduate school of engineering and science (University of the Ryukyus), and JICA-PEAC program for Afghanistan.

REFERENCES

[1] P. Aleotti, R. Chowdhury. (1999). "Landslide hazard assessment: Summary review and new perspectives," [Bulletin of Engineering Geology and the Environment]. 58(1), 21-44. Available: http://egsp.lyellcollection.org/content/18/1/43.abstract

[2] H. Vijith, K.N. Krishnakumar, G.S. Pradeep, M.V. Ninu Krishnan & G. Madhu , Georisk. (2013). "Assessment and Management of Risk for Engineered Systems and Geohazards Shallow landslide initiation susceptibility mapping by GIS-based weights-of-evidence analysis of multi-class spatial datasets a case study from the natural sloping terrain of Western Ghats, India, Georisk". [Assessment and Management of Risk for Engineered Systems and Geohazards]. Available: DOI:10.1080/17499518.2013.843437 and http://www.tandfonline.com/doi/abs/10.1080/17499518.2013.843437

[3] Xiwei Xu*. (2008). "Landslide Hazard Mapping Using GIS and Weight of Evidence Model in Qingshui River Watershed of Wenchuan Earthquake Struck Region". Available: DOI:10.1007/s12583-012-0236-7 and http://link.springer.com/article/10.1007/s12583-012-0236-7

[4] Geological Survey of Ireland. (2011). "Landslide Susceptibility Mapping: Literature Review and Findings". Available: http://www.gsi.ie/NR/rdonlyres/DCF40D67-DA2E-4282-AB09-8D31B4C0EBDE/0/DRAFTLiteratureReview.pdf

[5] DAVID K., KEEFER. (1984). "Landslides caused by earthquakes" [Geological Society of America Bulletin], v. 95, no. 4, p. 406-421. Available: Doi:10.1130/0016-7606(1984)95<406:LCBE>2.0.CO;2 and http://bulletin.geoscienceworld.org/content/95/4/406.full

[6] Available:http://poleshift.ning.com/m/blogpost?id=3863141%3ABlogPost%3A896468&maxDate=2014-07-21T13%3A06%3A55.569Z

[7] John F. Shroder, Megan Jensen Schettlera, Brandon J. Weihs b. (2011). "Loess failure in northeast Afghanistan". [Physics and Chemistry of the Earth]. 36, 1287–1293. Available: http://adsabs.harvard.edu/abs/2011PCE....36.1287S

[8] Russell L. Wheeler and Kenneth S. Rukstales. (2007). "Seismotectonic Map of Afghanistan and adjacent areas". [USAID]. Available: http://pubs.usgs.gov/of/2007/1104/downloads/pdf/of2007-1104.pdf

[9] UNDAC. Mission (1998, 10 February to 6 March). "A report following the earthquake of 4 February 1998 in Rustaq Afghanistan". Available: http://reliefweb.int/report/afghanistan/report-undac-mission-10-february-6-march-1998-following-earthquake-4-february

[10] Fausto Guzzetti. (2012). "Landslide inventory maps: New tools for an old problem". [Earth Science Reviews]. 112, 42–66. Available: http://www.sciencedirect.com/science/article/pii/S0012825212000128

[11] Williams C. J., S. S. Lee, R. A. Fisher, and L. H. Dickerman. (1999). "A Comparison of Statistical Methods for Prenatal Screening for Down syndrome". [Applied Stochastic Models in Business and Industry]. 15(2):89–101. Available: Doi: 10.1002/ (SICI) 1526-4025(199904/06)15:2<89: AID-ASMB366>3.0.CO; 2. And http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1526-4025(199904/06)15:2%3C89::AID-ASMB366%3E3.0.CO;2-K/abstract

Mohammad Kazem Naseri, Mohammad kazem naseri is a graduate of geology and mining from the Balkh University, Mazari-e-Sharif, Afghanistan. Presently, he is pursuing his master degree at the university of the Ryukyus, department of information engineering, Okinawa, japan. His area of interest include geology, GIS and remote sensing, landslide hazard mitigation and computer applications. He has the membership of The Institute of Electronics, Information and Communication Engineers (IEICE), since early 2015.

Dongshig Kang, Dongshik Kang is an Associate Professor at the Faculty of Information Engineering, University of the Ryukyus. His previous position was as an Assistant Professor at University of the Ryukyus (1999-2004). He received a Ph.D. in Engineering from the University of Osaka Prefecture. He is a member of the IEEJ, IEICE and GISA. His current research interests are neural networks, geographic information system, and intelligent systems.