Research Article
Single and Multi-source Methods for Reconstruction the Gaps in Landsat 7 ETM+ SLC-off Images

[1]Ghazali Sulong, [1, 2]Asmaa Sadiq and [3]LoayEdwar
[1]UTM-IRDA Digital Media Center (MaGIC-X), Faculty of Computing, Universiti Taknologi Malaysia, 81310, UTM Skudai Johor Takzin, Malaysia
[2]Department of Computer, College of Science, University of Al-Mustansiriyah, Baghdad, Iraq
[3]Department of Computer, College of Science, University of Baghdad, Baghdad, Iraq

Abstract: Since 2003, the Scan Line Corrector (SLC) instrument of the Enhanced Thematic Mapper plus (ETM+) sensor on board a Landsat 7 satellite has failed permanently causing regular gaps to appear in Landsat 7 images. This malfunction has been limited and hampered the scientific application of ETM+ data. Therefore, several methodologies and techniques have been conducted to reconstruct these gaps in order to expand the usability of the ETM+ SLC-off images. These methods can be classified as single source and multi-source methods. In this study two single source interpolation methods, mean and IDW are utilized to estimate the missing pixels value and the obtained results are compared with multi-source approach, LLHM. The results are assessed qualitatively and quantitatively using two statistical indicators RMSE and SE. The results indicated the superiority of LLHM on the single source interpolation methods.

Keywords: Gap filling, IDW, interpolation method, Landsat 7 ETM+, reconstruction, SLC-off image

INTRODUCTION

Since 1972, the Landsat series of satellites has gathered remotely sensed images from the earth's surface which provides the largest useful and dedicated data source of medium to high spatial resolution multispectral images. Landsat imagery provides a unique resource for scientists in the study of agriculture, forestry, education, mapping, regional planning, water management and urban development through monitoring human scale processes (Hu *et al*., 2011; Mohammdy *et al*., 2013; Sadiq *et al*., 2014).

The Landsat satellite sensors consist of the Landsat 1-5 Multispectral Scanners (MSS), the Landsat 5 Thematic Mapper (TM) and the Landsat 7 Enhanced Thematic Mapper plus (ETM+) (Chen *et al*., 2012). A small mirror in the optical path of the ETM+ instrument, Scan Line Corrector (SLC), which compensates for the forward motion of the satellite during acquisition suffered from a permanent failure on May 31, 2003 (Zhang *et al*., 2007; Boloorani *et al*., 2008a). As a result of SLC failure (SLC-off image), individual scan lines overlap each other alternately causing 22% of the pixels to be un-scanned and resulting in gaps as large as 14 pixels wide along the edges of the scene and gradually reducing near the center of the scene (Reza and Ali, 2008; Zhu *et al*., 2012a) Fortunately, about 80% of the pixels in ETM+

images being scanned, as well as the radiometric and geometric quality of these images has not been affected by SLC failure (Chen *et al*., 2010). Therefore, many methodologies and algorithms have been suggested for reconstructing the gaps of Landsat 7 SLC-off images. Generally, these methods can be categorized into single source and multi-source reconstruction methods.

In a single source method, the non-gapped areas are used to recover the gapped area in the SLC-off image itself (Boloorani *et al*., 2008b). The simplest case is the replacement of the missing pixel values with the average derived from the surrounding pixel values (Boloorani *et al*., 2008b). In addition, bilinear or bicubic interpolation algorithms were used as a direct interpolation approach to fill the gaps in SLC-off images (Hu *et al*., 2011). Thomas presented Inverse Distance Weight (IDW) method as a spatial single source interpolation method using existing known values within a moving window around the gaps pixels (Alexandridis *et al*., 2013). Zhang *et al*. (2007) presented geostatistical (kriging and co-kriging) gap filling technique which predicts the missing values in SLC-off image by exploiting spatial correlation information within the same area of the gaps. Also, Maxwel suggested the spectral interpolation method based on multi-scale segmentation models which was suitable for regional-scale studies, land cover and crop mapping (Zhu *et al*., 2012b).

Corresponding Author: Ghazali Sulong, UTM-IRDA Digital Media Center (MaGIC-X), Faculty of Computing, Universiti Taknologi Malaysia, 81310, UTM Skudai Johor Takzin, Malaysia

In multi-source method, the gap area is reconstructed using information derived from other images which are taken for the same scene rather than the SLC-off image itself (Desai and Ganatra, 2012). Two kinds of compositing methods, phase 1 and 2 product, were produced by the science office of the USGS (United States Geological Survey) Landsat project and NASA (National Aeronautics and Space Administration) Landsat project. In these products the SLC-off image was combined, as a primary image to be filled, with one or more previously acquired Landsat 7 (SLC-on image) or another SLC-off images, which are considered fill images, in order to restore the gaps locations in the malfunctioned image (Storey *et al.*, 2005). Several studies were derived depending on the principle of USGS/NASA product, Boloorani developed a multispectral projection transformation method based on the PCT to reconstruct the Landsat 7 ETM+ imagery using SLC-on imagery as auxiliary image (Boloorani *et al.*, 2008b) and J. Chen proposed simple and effective gap filling algorithm (NSPI) using auxiliary Landsat images to derive information related to neighboring pixels which are located close to gap pixels (Zhu *et al.*, 2012a). Also, several algorithms were proposed to fill the gaps using another auxiliary images rather than Landsat series images, which were detected at times as close as possible to the SLC-off image, Raza and Ali used IRS/ID LISS-III sensor product to reconstruct the malfunctioning of SLC-off image, while data derived from CBERS and EO-1/ALI sensors was used by Boloorani and Chen respectively to estimate the un-scanned pixel values (Chen *et al.*, 2012).

The goal of this research is to apply and validate a number of existing single source interpolation methods and USGS/NASA LLHM approach on Landsat 7 SLC-off images.

METHODOLOGY

Gap filling algorithm:
Single source method: In this category the gap locations in SLC-off images can reconstruct by applying either simple linear/nonlinear filters or using spatial interpolation methods on the malfunction landsat 7 image itself.

Mean filter: It is the simplest linear filter that gives equal weight to all pixels in an MxN neighborhood to replace the center pixel in the output image with the mean value extracted from this neighborhood. For a rectangular sub image window (R) of size MxN, the arithmetic mean or average values $\overline{X}$ of pixels $L (r, c)$ within this window is used to estimate the missing value (Ali and Mohammed, 2013):

$$\overline{X} = \frac{1}{MN} \sum_{(r,c) \in R} L (r, c) \tag{1}$$

Spatial interpolation methods: Generally, interpolation is a mathematical function aimed at estimating values at a location, where no measured values, fall between known values. Image interpolation is used in a variety of applications such as remote sensing and medical imaging (Azpurua and Ramos, 2010).

Spatial interpolation is based on Tobler geographic law where "everything is related to everything else, but near things are more related than distant things" (Ohashi and Torgo, 2012).

Therefore, in spatial interpolation the unknown value is calculated using a set of known values related to sample points located within an area. The values of these sample points and the distance between them and desired unknown value contribute to estimation of the required value. Generally, in spatial interpolation the set of known sample points is determined first. Then, using an interpolation method to estimate the unknown value, the size of sample points is limited to control the effect of these known values in the estimating process (Childs, 2004).

Inverse Distance Weight (IDW) interpolation method: It is a standard spatial interpolation method approved by geoscientists in geographic information science and based on spatial autocorrelation between the points in the study area (Lu and Wong, 2008; Jing and Wu, 2013). It is based on the assumption that measured nearest sample point has more influence or weighting in estimating the value of unknown center point. In other words, the significance of measured points can be controlled depending on the distance between these points and unknown data point, so the impact of these points decreases with the increase of the distance (Nusret and Dug, 2012; Childs, 2004).This interpolation approach is suitable in areas with regularly distributed points rather than cluster areas where the points are distributed in an irregular manner (Azpurua and Ramos, 2010).

The mathematical equation used to calculate the interpolated values Y at a location X using a set of surrounding measured values $Y_i = Y (X_i)$ where i = 0, 1, 2,.., M is:

$$IDW = Y(X) = \sum_{i=0}^{M} \frac{W_i(X)Y_i}{\sum_{j=0}^{M} W_j} \tag{2}$$

where,
M = To the number of measured sample points within the neighborhood
W_i = A simple Shepard weighting function, $W_i(X) = \frac{1}{d (X,X_i)^p}$
d = The distance between unknown point X and measured point X_i
p = The positive power factor that indicates the rate of decrease of the weights as distance increases (Jing and Wu, 2013; Peralvo, 2003)
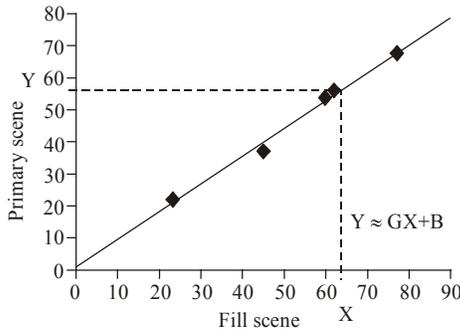
Fig. 1: An assumed linear relationship between primary and fill scenes, from Scaramuzz *et al*. (2004)

IDW equation indicates that smaller weights will be assigned to the measured points which are further from an unknown point *X* and vice versa. In addition, a larger value of *p* implies the assignment of larger weights to the nearest measured points and at the same time decreasing weights given to the furthest points (Lu and Wong, 2008).

Multi source method: In multi-source methods, preliminary studies in reconstructing the gap locations of SLC-off image were performed by using another image of the Landsat 7 ETM+ before or after failure for the same scene (SLC-on or SLC-off) in order to estimate the values in gaps locations.

LLHM technique: After SCL failure, research scientists of the United States Geological Surv (USGS) National Centre for Earth Resources Observation and Science (EROS) presented gap filling products consisting of phase 1 and 2 using multi source imagery (more than one image with the same path and row). Both phases have been designed to combine SLC-off image with one or more SLC-off and/or SLC-on image using histogram based compositing algorithms to find a linear transformation between two input images as shown in Fig. 1. Phase 1 consists of Global Histogram Matching (GHM) and Local Linear Histogram Matching (LLHM) techniques which are characterized by their ease and speed of implementation and it is performed using SLC-off scene as primary image and one SLC-on scene as fill image (Scaramuzz *et al*., 2004; Zhang *et al*., 2007).

In LLHM technique, after the gaps and the valid pixels are found in primary image, the linear transform between the primary and fill images was performed using a moving window over all un-scanned pixels in the scene. After calculating the common pixels between primary and fill images within the moving window the corrective gain and bias of input data in both target and fill image were calculated to produce the pixel values of the primary image. Mean and standard deviation are used rather than using a computationally expensive

linear fit to calculate the gain and bias. This was performed using the following equations:

$$G = gain = \frac{\sigma_p}{\sigma_f} \tag{3}$$

$$B = bias = \mu_p - \mu_f * gain \tag{4}$$

where, μ_p, σ_p, μ_f, σ_f are the mean and standard deviation of the primary and the fill image window respectively calculated from:

$$\mu_p = \frac{1}{N}\sum_{i=1}^{N} primary\ DN_i \tag{5}$$

$$\sigma_p = \frac{1}{(N-1)}\sum_{i=1}^{N}(primary\ DN_i - \mu_p)^2 \tag{6}$$

$$\mu_f = \frac{1}{N}\sum_{i=1}^{N} fill\ DN_i \tag{7}$$

$$\sigma_f = \frac{1}{(N-1)}\sum_{i=1}^{N}(fill\ DN_i - \mu_f)^2 \tag{8}$$

Then the new value (*Y*) is used to fill the gap in a primary scene obtained from calculated gain and bias and the corresponding fill image pixel value (*X*) using the following equation (Scaramuzz *et al*., 2004; USGS/NASA, 2004):

$$Y = X* G+B \tag{9}$$

A number of conditions are required for both the primary and fill image to prevent limitations and poor results of LLHM technique. These requirements are minimal snow cover and clouds, minimal data separation and low temporal variability between the primary and the fill images (USGS/NASA, 2004; Zhu *et al*., 2012b).

EXPERIMENTAL RESULTS

The region located on Path 39 and Row 35 in World Reference System 2 (WRS-2) is selected for simulated experimentation. The L1G product acquired from the USGS website (http://glovis.usgs.gov/) is used as a dataset. Three true color bands (R = band 3, G = band 2 and B = band 1) with 300×300 pixel sub images are used and stored as the original 8-Byte integer value. Both simulated and actual ETM+SLC-off images with moderate pixel width (5-7 gap wide) are employed to test and compare the performance of two single reconstruction approach, mean and IDW with LLHM method.

Experimental results on simulated SLC-off images: For simulation test, subimage is selected with minimum cloud cover and good quality for precise results. The segment image selected is dominated by land cover change like coastline of a river surrounded by sedimentary area and rocky lands as shown in Fig. 2.
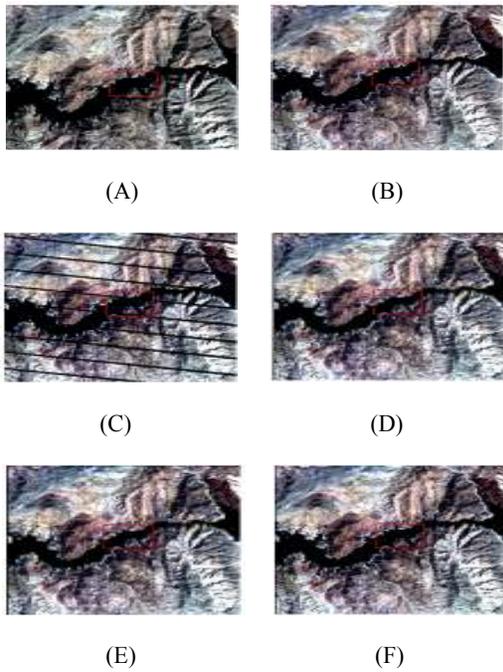
(A)   (B)

(C)   (D)

(E)   (F)

Fig. 2: Red-green-blue composites of Landsat 7 SLC-on image for the simulation captured in 2003; (A): Target image in May 26; (B): SLC-on fill image of March 9. Panel; (C): The simulated SLC-off based on image from A. PANELS; (D-F): The results of mean, IDW and LLHM approaches respectively
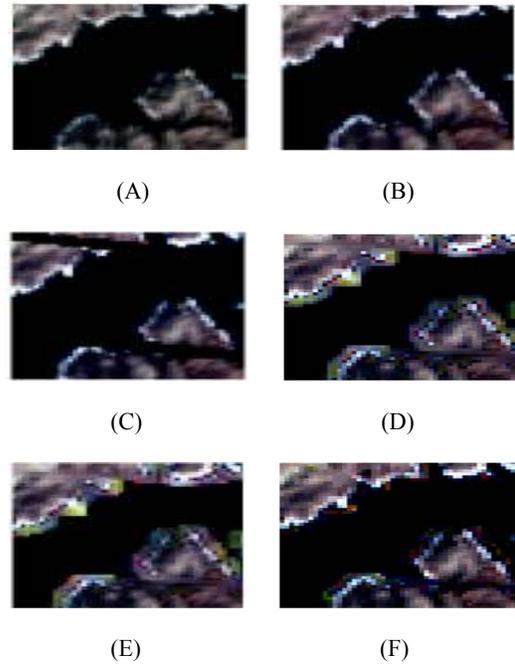

(A)   (B)

(C)   (D)

(E)   (F)

Fig. 3: Zoomed regions cropped from Fig. 2; (A): Target SLC-on image; (B): The fill SLC-on image; (C) The simulated SLC-off of the target image. Results of; (D): mean; (E): IDW; (F): LLHM approaches

To assess the gap filling algorithms, both the qualitative and quantitative evaluations are used. Visual observation is used to evaluate the results qualitatively to check whether all the gaps are filled or if there are still strips appearing in the reconstructed images, while statistical indices are used to evaluate the results quantitatively by comparing the filled image with actual image data (Chen *et al*., 2011, 2010). In our research, two statistical indicators are used to evaluate the accuracy of the obtained results, Root Mean Square Error (RMSE) and Systematic Error (SE) which are used to assess the differences between the estimated value obtained from the gap filling algorithm and the actual observed value using the following equations:

$$RMSE_f = \sqrt{\frac{\sum_{i=1}^{N}(\bar{k_i}-k_i)^2}{N}} \qquad (10)$$

$$SE_f = \frac{\sum_{i=1}^{N}(\bar{k_i}-k_i)^2}{N} \qquad (11)$$

where, $\bar{k_i}$, k_i are the estimated and actual values of the ith gap pixel in the reconstructed and fill images respectively and N is the number of un-scanned pixels. Larger RMSE and SE values indicate a larger estimation error (Chen *et al*., 2010).

Figure 2A and B illustrates two Landsat 7 SLC-on images acquired on May 26 and March 9, 2003,

respectively. In addition, Fig. 2C shows the simulated image of Fig. 2A, the simulated image is created using an actual or real SLC-off image, where the zero values of the gap locations replaced the pixel values in the same location of the target image. Subsequently Fig. 2C is considered as a target image to be filled while Fig. 2B is used as an auxiliary fill image in multi-source reconstruction approach. Also the actual or real SLC-on image in (B) is used to assess the performance of the single and multi-source reconstruction approaches by comparing the filled images with the real one.

Figure 2D and E shows the results of applying the single source reconstruction methods (mean and IDW respectively), which are easy and simple to implement. It is obvious that in both the mean and IDW approaches that in spite of all the gaps being filled, results can be obviously inaccurate, such as the appearance of a clear boundary between the recovered pixels and the target image. This shortcoming is because the fill value has been derived from the surrounding pixel value rather than calculating the reflectance of the land. Figure 2F shows the LLHM multi source algorithm result and as shown, most land types are sufficiently restored in the filled image and the boundaries between the recover pixels and target image are disappearing. However, there are performance limitations around land cover changes such as coastlines. For more details Fig. 3 shows the zoomed region cropped from the results of Fig. 2 (zoomed region specified via a fixing red

Table 1: The results of the quantitative assessment of the reconstruction algorithms

	Mean filter		IDW filter		LLHM technique	
	RMSE	S.E.	RMSE	S.E.	RMSE	S.E.
B1	0.082957	0.006881	0.078557	0.004592	0.043655	0.0019057
B2	0.105186	0.011061	0.116186	0.013499	0.037099	0.0013760
B3	0.113186	0.012811	0.137186	0.004253	0.049880	0.0024880

S.E.: Standard error

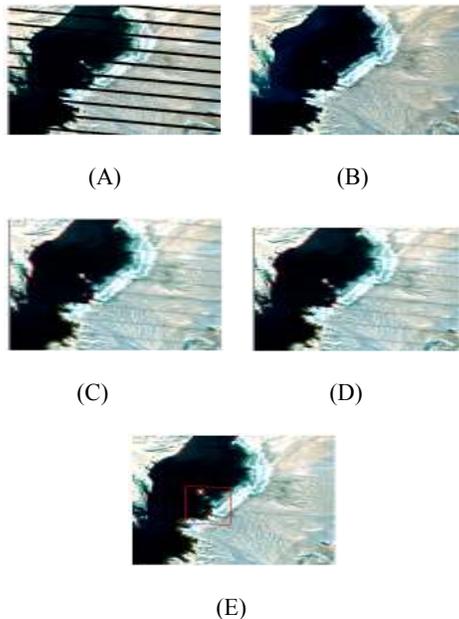(A)   (B)

(C)   (D)

(E)

Fig. 4: Experiment results on actual Landsat 7 SLC-off image (taken in 2003); (A): The target image (September 17); (B): Fill image (April 26). The obtained result; (C): From mean method; (D): From IDW method; (E): From LLHM method

window). The limitation of the LLHM technique is in the case of radical differences in target radiance because of the changes in terrestrial cover. This limitation is due to different time acquisition or with heterogeneous landscape, where; the size of the surface objects in reconstructing area are smaller than the local moving window.

Table 1 shows the quantitative assessment of the reconstruction algorithm results for each band using RMSE and SE statistical indices. The statistical results indicate that in single source methods the results of mean and IDW methods are similar to larger statistical values due to its inaccurate performance. The results also show the superiority of the LLHM approach over single source methods (lower values of RMSE and SE) in all bands and this is identical to qualitative results.

Experimental results on actual SLC-off images: Further performance evaluation of the single and multi-source reconstruction methods is carried out by applying them on real or actual Landsat 7 SLC-off images. Figure 4 illustrates the real SLC-off sub image selected from the same region in simulated testing with

a moderate pixel width (7 pixels wide) and following the previous simulated case study, the actual SLC-off image acquired on September 17, 2003 as shown in Fig. 4A is used as the target image to be filled while the SLC-on image taken on April 26, 2003 shown in Fig. 4B is used as fill image with LLHM approach. Figure 4C, D and E presents the results of the mean, IDW and LLHM methods, respectively. Although the LLHM method is found to achieve a good visual clarity where most of the gaps are recovered with superior quality of the filled image as compared to the mean and IDW method which displays shadow in the recovered areas, there are obvious artifacts around small and sharp edges (as shown in red box) where the histogram of the data is very narrow.

CONCLUSION

In spite of the malfunction of the SLC in Landsat 7 satellite images, the radiometric and geometric quality of ETM+ sensor had not been affected, therefore Landsat 7 imagery has been broadly used in many regional and global studies. Several methodologies and studies are developed in reconstructing the SLC-off images. In this study the results of two single source algorithms are compared with a USGS LLHM approach using statistical indices in addition to qualitative assessment.

From the results derived from simulated and actual study area, in general, the information of the spatial pattern is not used in single source interpolation algorithms; therefore these methods may not be appropriate in reconstructing the gaps in Landsat 7 SLC-off image. In addition, LLHM approach examines the pixels within a moving window; therefore it can yield suitable results over scenes with invariant terrain such as deserts and rocky areas (homogeneous landscape). However, the evolution of different patterns in different land cover is ignored in the heterogeneous landscapes.

Therefore, in gap filling algorithms, the spatial constraint (the consistent of the desired pixel with surrounding pixel values) and temporal constraint (the information from auxiliary images) can be used to estimate the values in un-scanned locations.

ACKNOWLEDGMENT

REFERENCES

Alexandridis, T.K., I. Cherif, C. Kalogeropoulos, S. Monachou, K. Eskridge and N. Silleos, 2013. Rapid error assessment for quantitative estimations from Landsat 7 gap-filled images. Remote Sens. Lett., 4(9): 920-928.

Ali, S.M. and M.J. Mohammed, 2013. Gap-filling restoration methods for ETM+ sensor images. Iraqi J. Sci., 54(1): 206-214.

Azpurua, M. and K. Ramos, 2010. A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. Prog. Electromagn. Res. M, 14: 135-145.

Boloorani, A.D., S. Erasmi and M. Kappas, 2008a. Multi-source remotely sensed data combination: Projection transformation gap-fill procedure. Sensors, 8(7): 4429-4440.

Boloorani, A.D., S. Erasmi and M. Kappas, 2008b. Multi-source image reconstruction: Exploitation of EO-1/ALI in landsat-7/ETM+ SLC-off gap filling. Proceeding of the IS&SPIE's 20th Annual Symposium, Electronic Imaging Science and Technology. San Jose, California, USA, January 27-31, 6812: 681219-1 to 681219-12.

Chen, F., L. Tang and Q. Qiu, 2010. Exploitation of CBERS-02B as auxiliary data in recovering the landsat 7 ETM+ SLC-off image. Proceeding of the 18th International Conference on Geoinformatics. Beijing, pp: 1-6.

Chen, F., X. Zhao and H. Ye, 2012. Making Use of the Landsat 7 SLC-off ETM+ Image through Different Recovering Approaches. In: Karakehayov, Z. (Eds.), Data Acquisition Applications. Chapter 13, ISBN: 978-953-51-0713-2, In Tech, DOI: 10.5772/48535.

Chen, J., X. Zhu, J.E. Vogelmann, F. Gao and S. Jin, 2011. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. Remote Sens. Environ., 115(4): 1053-1064.

Childs, C., 2004. Interpolating Surfaces in ArcGIS Spatial Analyst. ESRI Educational Services. Retrieved from: http://www.esri.Com/news/arcuser/0704/files/inter polating.pdf.

Desai, M. and A. Ganatra, 2012. Survey on gap filling in satellite images and inpainting algorithm. Int. J. Comput. Theor. Eng., 4(3): 341-345.

Hu, W., M. Li, Y. Liu, Q. Huang and K. Mao, 2011. A new method of restoring ETM + SLC-off images based on multi-temporal images. Proceeding of the 19th International Conference on Geoinformatics. Shanghai, pp: 1-4.

Jing, M. and J. Wu, 2013. Fast image interpolation using directional inverse distance weighting for real-time applications. Opt. Commun., 286: 111-116.

Lu, G.Y. and D.W. Wong, 2008. An adaptive inverse-distance weighting spatial interpolation technique. Comput. Geosci., 34(9): 1044-1055.

Mohammdy, M., H.R. Moradi, H. Zeinivand, A.J.A.M. Temme, H.R. Pourghasemi and H. Alizadeh, 2013. Validating gap-filling of landsat ETM+ satellite images in the Golestan Province, Iran. Arab. J. Geosci., 7(9): 3633-3638.

Nusret, D. and S. Dug, 2012. Applying the inverse distance weighting and Kriging methods of the spatial interpolation on the mapping the annual Precipitationin in Bosnia and Herzegovina. Proceeding of 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting. Leipzig, Germany.

Ohashi, O. and L. Torgo, 2012. Spatial interpolation using multiple regression. Proceeding of the IEEE 12th International Conference on Data Mining, pp: 1044-1049, DOI: 10.1109/ICDM.2012.48.

Peralvo, M., 2003. Influence of DEM interpolation methods in Drainage Analysis. Proceeding of GIS Hydro 2004 Preconference Seminar Website, GIS in Water Resources Term Projects. Retrieved from: http://www.crwr.utexas.edu/gis/gishydro04/. (Accessed on: July 17, 2013)

Reza, M.M. and S.N. Ali, 2008. Using IRS products to recover 7ETM + defective images. Am. J. Appl. Sci., 5(6): 618-625.

Sadiq, A.J., G. Sulong and L.E. George, 2014. Survey on gap filling algorithms in landsat 7 ETM + images. J. Theor. Appl. Inform. Technol., 63(1): 136-146.

Scaramuzz, P., E. Micijevic and G. Chander, 2004. SLC Gap-filled Products Phase One Methodology. Retrieved from: http://landsat.usgs.gov/documents/SLC_Gap_Fill_ Methodology.pdf. (Accessed on: March, 2013)

Storey, J., P.S. Engineer and S. Falls, 2005. Landsat 7 scan line corrector-off gap-filled product gap-filled product development. Proceeding of the 6th Pecora Conference American Society of Photogrammetry and Remote Sensing. Bethesda, Maryland, Unpaginated CD-ROM, October 23-27, pp: 1-13.

USGS/NASA, 2004. SLC-off Gap-filled Products Gap-fill Algorithm Methodology: Phase 2. Gap-fill Algorithm. Retrieved from: www.ga.gov.au/servlet/BigObjFileManager?bigobj id=GA4861. (Accessed on: October, 2013)

Zhang, C., W. Li and D. Travis, 2007. Gapsfill of SLC-off landsat ETM+ satellite image using a geostatistical approach. Int. J. Remote Sens., 28(22): 5103-5122.

Zhu, X., F. Gao, D. Liu and J. Chen, 2012a. A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images. IEEE Geosci. Remote S., 9(3): 521-525.

Zhu, X., D. Liu and J. Chen, 2012b. A new geostatistical approach for filling gaps in landsat ETM+ SLC-off images. Remote Sens. Environ., 124: 49-60.