A Statistic Correlation Analysis Algorithm Between Land Surface Temperature and Vegetation Index

Hyung Moo Kim*, Beob Kyun Kim*, and Kang Soo You**

Abstract: As long as the effective contributions of satellite images in the continuous monitoring of the wide area and long range of time period, Landsat TM and Landsat ETM+ satellite images are surveyed. After quantization and classification of the deviations between TM and ETM+ images based on approved thresholds such as gains and biases or offsets, a correlation analysis method for the compared calibration is suggested in this paper.

Four time points of raster data for 15 years of the highest group of land surface temperature and the lowest group of vegetation of the Kunsan city Chollabuk_do Korea located beneath the Yellow sea coast, are observed and analyzed their correlations for the change detection of urban land cover.

This experiment based on proposed algorithm detected strong and proportional correlation relationship between the highest group of land surface temperature and the lowest group of vegetation index which exceeded R=(+)0.9478, so the proposed Correlation Analysis Model between the highest group of land surface temperature and the lowest group of vegetation index will be able to give proof an effective suitability to the land cover change detection and monitoring.

Keywords: LST, NDVI, Correlation Analysis, Landsat ETM+

1. Introduction

This paper suggests a correlation analysis model between the NDVI lowest area and the highest land surface temperature area which uses Landsat-5 TM(thematic mapper) with Landsat-7 ETM+(enhanced thematic mapper plus) satellite image in order to extract change pattern for the change detection and the spatio-temporal change patterns of urban area surface temperature, suburban area land cover and vegetation.

The experiment results of applying the proposed model into change detection in Chollapuk_do Kunsan city area shows that correlation analysis model between the NDVI lowest area and the highest land surface temperature area which uses Landsat-5 TM (thematic mapper) with Landsat-7 ETM+ (enhanced thematic mapper plus) satellite image was very efficient in land cover change detection.

Markham and Baker presented post-validation coefficient index value for the substitution of Planck's function at each sensor quality, which was needed to apply into the equation conversion between DN (Digital Number) and spectral radiation of black body when to calculate land surface temperature (LST) using Landsat MSS and Landsat TM[3].

Also, Chander and Markham proposed a new revision process and dynamic segment range of specific constant like gain or offset in each band of radiation correction which reflects the changes of Markham and Baker's equation relationship condition[1].

NASA suggested the approval dynamic scope of gain and offset in Landsat ETM+ which had not been proposed by Chander and Markham[5].

Melesse extracted land cover specific coefficient which was calculated and analyzed from NDVI (Normalized vegetation index) then, explained rainfall outflow which reflects the flow changes for 28 years due to the water which was melted at snowed mountains where land and the cities were located beneath[4].

This paper is based on NASA method [2] and applies Melesse model to Kunsan city area then extract a continuous distribution pattern between LST and NDVI using ERDAS IMAGINE 8.7.

2. Study Area and Data

2.1 Study Area

Industrial jar is located in the west bound of the study area, Kunsan city, and the agricultural plain is located in

Fig. 1. Study area: Kunsan city Chollabuk_do, Korea

Manuscript received October 4, 2005; accepted November 14, 2005.

* Department of Computer Engineering, Chonbuk National University, Jeonju, Korea ({kimhyungmoo, kyun}@chonbuk.ac.kr)
** School of liberal arts Jeonju University, Jeonju, Korea (gsyou@jj.ac.kr)

ISSN 1738-8899 ⓒ 2005 KIPS

the south bound. Saemangeum reclamation district of 401.00 ㎢ is located between the south bound of Kunsan city and Yellow Sea. North and east bound of Kunsan is relatively forest area.

2.2 Data

From band 1 to band 5 and band 7 have same sensor characteristics among Landsat TM and ETM+. Band 6 of ETM+ has two sub bands named low and high. Only ETM+ has panchromatic band 8.

Table 1. Landsat 7 TM/ETM+ band specifications

Feature \ Band	Band1	Band2	Band3	Band4	Band5	Band 7	Band8	Band 6	
	TM/ETM+						ETM+	TM/ETM+	
Name	Visible			NIR	MIR	SIR	Panchromatic	TIR	
	B	G	R					TM	ETM+
Spectral Resolution (㎛)	0.45-0.52	0.53-0.61	0.63-0.69	0.78-0.90	1.55-1.75	2.09-2.35	0.52-0.90	10.40-12.50	
Spatial Resolution (m*m)	30*30	30*30	30*30	30*30	30*30	30*30	15*15	120*120	60*60
Temporal Resolution (day)	16	16	16	16	16	16	16	16	

The specification detail of 4 time points images are described at table 1. 4 images are selected for representing each 4 unique seasons spring, fall, summer and winter, respectively. Acquired time of each image is ten to twelve o'clock a.m.

Table 2. Landsat TM/ETM+ image pre-processing

No.	Year	Sensor	Pre-processing
1	1987.04.18	Landsat5TM	Geo-rectification
2	1995.10.17	Landsat5TM	Geo-rectification
3	1999.05.21	Landsat5TM	Geo-rectification
4	2002.02.14	Landsat7ETM+	Geo-rectification

3. Image Pre-processing

3.1 Land cover Pre-processing

The satellite images used in this study, are distributed by ETRI and pre-processed by their image pre-processing component, then extracted the region of interest according to the 1:25,000 scale administrative boundary layer of National Geographic Intelligence Agency of Korea. After the geo-rectification, all images are classified by non-supervised ISODATA clustering algorithm which is built in the ERDAS IMAGINE 8.7.

3.2 NDVI Pre-processing

NDVI is the post-calibrated index of green intensity on the vegetation cover[6]. The principal of NDVI is that the reflexes rates are differ at the NIR and band 3, so these differences can produce an image which represents the conditions of green plant. This equation is denoted as following formula (1)[4].

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

Following Fig. 2 shows the algorithm to get the land surface temperature, NDVI specific coefficient ε (land cover specific coefficient ε), then, analyze the correlation between LST (land surface temperature) and NDVI of continuously changing land cover.

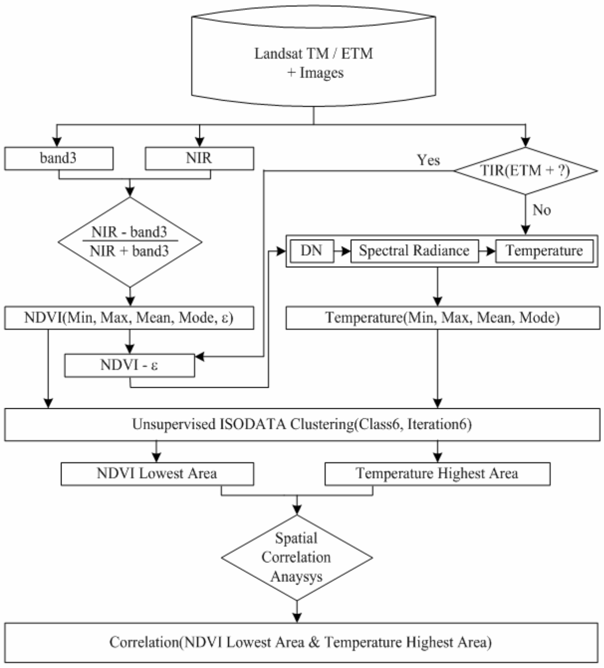


Fig. 2. Flow chart of correlation analysis algorithm

In the Fig. 2 when the NDVI of ETM+ images are greater than zero, we extract land cover specific coefficient ε, the substitute it as constant K_1 to get the LST, otherwise we get the LST directly from TM or ETM+ images. The rest of the proposed algorithm is correlation analysis between LST highest area and NDVI lowest area after ISODATA clustering.

3.3 LST Pre-processing

3.3.1 NASA equation between spectral radiance and temperature

Based on NASA model, this study calibrates surface temperature from the DN (Digital Number) which represents the absolute radiation of land cover. First of all, when each DN of TM and ETM+ images is given, we can subtract the spectral radiance L_λ ($L_{\lambda\text{-TM}}$, $L_{\lambda\text{-ETM+}}$) using the official NASA approval ranges $LMIN_\lambda$ and $LMAX_\lambda$ in following formula (2)[1] [4] [5].

$$L_\lambda = \frac{LMAX_\lambda - LMIN_\lambda}{(Q_{\mathrm{cal\,max}} - Q_{\mathrm{cal\,min}}) \times (Q_{\mathrm{cal}} - Q_{\mathrm{cal\,min}})} + LMIN_\lambda \quad (2)$$

a. Spectral radiance estimation in Landsat TM

By the above formula (2), in case of the least post-calibration value Q_{calmin} (DN) is equal to zero, the $L_{\lambda\text{-TM}}$ is able to be calculated by following linear expression formula (3)[1] [5].

$$L_{\lambda-TM} = G_{rescale} \times Q_{cal} + B_{rescale} \quad (3)$$

Table 3 is the example of official dynamic range of post-calibration scale values.

Table 3. Landsat TM spectral radiance, Post-calibration $G_{rescale}$, $B_{rescale}$ dynamic ranges

Band	$LMIN_\lambda$	$LMAX_\lambda$	$G_{rescale}$	$B_{rescale}$
6	1.2378	15.303	0.055158	1.2378

b. Spectral radiance estimation in Landsat ETM+

In the above formula (2), LPGS (ESO Data Gateway) uses 1 as the least post-calibration value Q_{calmin}, while NLAPS (Earth Explorer) uses 0. Based on this folicy, we can get the absolute spectral radiance by the following formula (4)[5].

$$L_{\lambda-ETM} = \text{"}gain\text{"} \times Q_{cal} + \text{"}offset\text{"} \quad (4)$$

There is no need to rectify the spectral radiance value in ETM+ because the two subands in ETM+ band 6, named Low gain 6 (1) and High gain 6 (2) are separated always.

Table 4. Landsat ETM+ spectral radiance LMINλ and LMAXλ "offset" "gain" ranges

Band	Low Gain		High Gain	
	$LMIN_\lambda$	$LMAX_\lambda$	$LMIN_\lambda$	$LMAX_\lambda$
6	0.0	17.04	3.2	12.65

3.3.2 Landsat TM/ETM+ Temperature

a. NASA model to extract temperature from TM/ETM+ images

As discussed in the above formulas (1) to (4), there is the relationship between the spectral radiance value L_λ ($L_{\lambda\text{-TM}}$, $L_{\lambda\text{-ETM+}}$) and the absolute temperature °K(Kelvin). This can be denoted like following formula (5)[4] [5].

$$T(°K) = \frac{K_2}{ln((K_1 / L_\lambda)+1)} \quad (5)$$

In the formula (5), DN means raster pixel value, L_λ ($L_{\lambda\text{-TM}}$, $L_{\lambda\text{-ETM+}}$) means the energy strength of solar light, and T means the absolute temperature of the land surface. K_1 is the post-calibration constant of spectral radiance, and k_2 is the post-calibration constant of absolute temperature.

Table 5. Landsat 5/7, TM/ETM+ thermal band calibration constants

sensors / constants	K_1	K_2
Landsat-5 TM	607.76	1260.56
Landsat-7 ETM+	666.09	1282.71

b. Melesse model to extract temperature from NDVI coefficient

Melesse model uses NDVI coefficient, the extract temperature from the same Planck function presented at formula (6)[4].

$$T_{NDVI}(°K) = \frac{K_2}{ln((\varepsilon \times K_1 / L_\lambda)+1)} \quad (6)$$

In the formula (6), T means absolute temperature of land surface, L_λ is calibrated by formula (2) with options like $Q_{calmin} = 1$, $Q_{calmax} = 255$, $L_{max} = 17.04$ [W/(m^{2}sr^1 μ m^1)], and $L_{min} = 0$ [W(m^{2}sr^1 μ m^1)]. ε means band specific coefficient and is came from NDVI which is calculated by formula (1). In case of NDVI > 0, ε is 1.009 + 0.047 * (ln NDVI), otherwise ε takes constant 1. K_1 means spectral radiance post-calibration constant 666.09 (Landsat-7 ETM+), and K_2 means temperature post-calibration constant 1282. 71K (Landsat-7 ETM+).

4. Conclusion and Discussion

4.1 Experiment Result

Land cover specific coefficient εs that canbe substitute spectral radiance post-calibration constant K_1 of ETM+ images are presented at table 6. NDVI-means in the study area are in the range of positive until 1995, then are in the range of negative from 1999. In the year 2002, since the image is an ETM+ and the NDVI-mean indicates -0.014 and the NDVI-mode indicates -0.00044, NDVI- ε should be 1. However we separate 2002.02.14 as 2002.02.14-a and 2002.02.14-b.

Table 6. NDVI ranges and NDVI-εfor estimation of land surface temperature

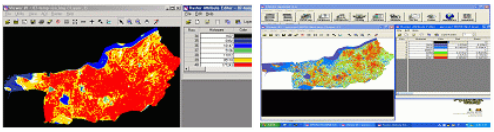
NDVIs / Year	NDVI-Min	Max	Mode	Mean	ε
1987.04.18	-0.5	0.543	-0.00095	0.0013	0.6966
1995.10.17	-0.705	0.688	-0.00024	0.0533	0.8712
1999.05.21	-0.571	0.591	-0.00146	-0.0003	1
2002.02.14-a	-0.508	0.496	-0.00044	-0.014	1
2002.02.14-b	0.113	0.113	0.113	0.113	0.9065

In case of 2002.02.14-b , land cover specific coefficient ε* K_1 substitute K_1. As the result of this substitution, temperature increases from +6.55 to +11.817, and the deviation increases -1.452 to -6.717. These means that 2002.02.14-b method is worse than 2002.02.14-a.

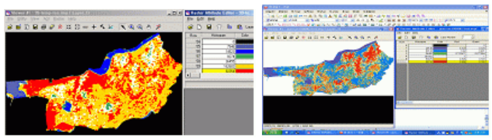
Table 7. Observed temperature and estimated land surface temperature(unit: °C)

Temp.s / Year	Observed	Landsat TM/ETM+ LST	Dev.
2002.02.14-a	5.1	6.552	-1.452
Avr.	17.3	16.45	0.85
2002.02.14-b	5.1	11.817	-6.717
Avr.	17.30	17.77	-0.47

This paper applied the method 2002.02.14-a to get the land surface temperature from satellite images and the method of 2002.02.14-b was not applied.


(a) 1987.04.18 LST (b) 1987.04.18 NDVI


(c) 1995.10.17 LST (d) 1995.10.17 NDVI


(e) 1999.05.21 LST (f) 1999.05.21 NDVI

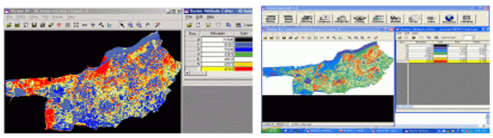

(g) 2002.02.14 LST (h) 2002.02.14 NDVI

Fig. 3. Land cover clustering of LST and NDVI

Four time points of Kunsan city area images representing four unique seasons are selected and pre-processed, then classified by the non-supervised ISODATA clustering algorithm. (b), (d), (f) and (h) of Fig. 3 show that NDVI lowest areas are changing to the changes of compared LST.

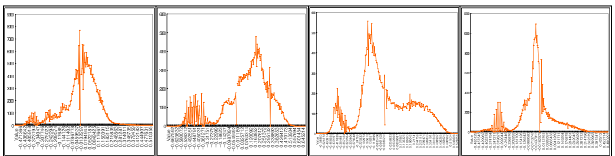

(a)1987.04.18 (b)1995.10.17 (c)1999.05.21 (d) 2002.02.14

Fig. 4. Kunsan city NDVI histogram

When it sees Fig. 4, the left half from NDVI zero is growing sharpen more as the time passing. The result of experiment that analyze the correlation between this NDVI lowest area and the LST highest area is presented at table 8.

Table 8. Correlation coefficient between the highest LST and the lowest NDVI

Variables / Year	LST Highest (㎢)	NDVI Lowest (㎢)
LST Highest	1	
NDVI Lowest	R=(+)0.947822772	1

In the Kunsan city area, LST highest area changes within the range of 42.40 ㎢ -159.51 ㎢, the median is 75.02㎢, the standard error is 25.26㎢, and the SD is 50.53 ㎢. Also, NDVI lowest area changes within the range of 21.60㎢ – 27.37 ㎢, the median is 24.27㎢, the standard error is 1.25㎢, and the SD is 2.50㎢.

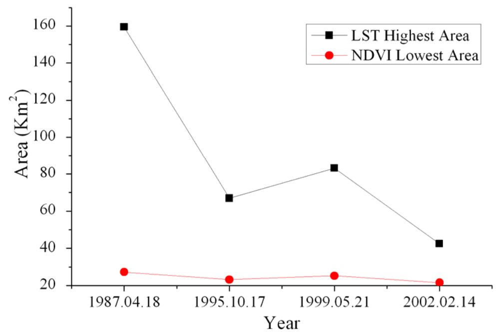


Fig. 6. Correlation scattergram between the highest LST and the lowest NDVI

4.2 Evaluation

The result of experiment shows that the max value of LST highest area is 112.1 ㎢ in 1997, the min value is 67.9 ㎢ in 1995, the max value of NDVI lowest area is 27.3 ㎢ in 1987, and the min value is 21.6 ㎢ in 2002. The overall pearson's correlation coefficient is r = (+) 0.9478 between LST highest area and NDVI lowest area. The result corresponds to the conventional researches that there is negative correlation between LST and NDVI. Furthermore our result shows that the LST highest area and NDVI lowest area are the independent variables in the study of surface change detection.

5. Conclusion

In order to find out the efficient method to detect the changes of cities, this study observed land surface temperature of land cover and vegetation pattern after designing a correlation analysis between LST highest area and NDVI lowest area. This paper defined the process to convert the land surface temperature from satellite images, those are varying Landsat7 TM/ETM+ respectively. Especially, in case of ETM+ we adopted Mellesse's NDVI specific LST calculation coefficient then we applied them to Chollabuk_do Kunsan city land cover change detection. This experiment shows the result that there is the overall pearson's correlation coefficient is r = (+) 0.9478. With this result we suggest that the raster based LST highest area-NDVI lowest area model is very efficient in the field of land cover detection.

References

[1] Chander, Gyanesh and Brian L. Markham. 2003. Revised Landsat-5 TM Radiometric Calibration Procedures and Post-calibration Dynamic Ranges, IEEE Transactions on geoscience and remote sensing, 41(11):2674-2677.

[2] Fisher, Jeremy I., John F., Mustard. 2004. High spatial resolution sea surface climatology from Landsat thermal infrared data, Remote Sensing of Environment 90:293-307.

[3] Markham, Brian L., and John L. Baker. 1986. Landsat MSS and TM Post-Calibration Dynamic Ranges, Atmospheric Reflectance and At-Satellite Temperatures, Laboratory for Terrestrial Physics- NASA/Goddard Space Flight Center, Greenbelt, MD. 20771, pp.3-7.

[4] Melesse, Assefa M. 2004. Spatiotemporal dynamics of land surface parameters in the Red river of the north basin, Physics and Chemistry of the Earth 29:795-810.

[5] NASA. 2004. Landsat Project Science Office. Landsat 7 Science Data Users Handbook. Chapt.11- Data Products, http://ltpwww.gsfc.nasa.gov/IAS/handbook/handbook_htmls/chapter11/chapter11.html:11.1-11.4.

[6] Rouse, J. W., R. H. Haas, J. A. Schell. 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Proc., Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, NASA SP-351, pp.3010-3017.

[7] Suga, Y., Ohno, H. Ogawa, K. Ohno, K. Yamada. 2003. Detection of surface temperature from Landsat-7/ETM+. Adv. Space Res. 32(11):2235-2240.

Hyung Moo Kim
1987 BA Seoul National University
1990 MA Seoul National University
1997 MS New York, Long Island University
2006 Ph.D. Chonbuk National University
Affiliation : Department of Computer Engineering, Chonbuk National University
Research Area : Object-orient Modeling, GIS, RS, Software Engineering, Telematics, Ubiquitos

Kang Soo You
1991 BA Chonbuk National University
1994 MA Chonbuk National University
2005 Ph.D. Chonbuk National University
1996 ~ : Visiting Professor, School of Liberal arts, Jeonju University
Research Area : Signal Processing, JPEG/MPEG, Lossless Image Compression, GIS

Beob Kyun Kim
1994 BA Chonbuk National University
1997 MA Chonbuk National University
2005 Ph.D. Chonbuk National University
2005 ~ : Researcher, RCIT, Chonbuk National University
Research Area : Distributed & Parallel Processing, Grid Computing, GIS