ISPRS Journal of Photogrammetry and Remote Sensing 96 (2014) 20–27

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Geostatistical estimation of signal-to-noise ratios for spectral vegetation indices

Lei Ji [a,*], Li Zhang [b,c], Jennifer Rover [d], Bruce K. Wylie [d], Xuexia Chen [e]

[a] ASRC InuTeq, Contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198-0001, USA
[b] Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
[c] State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
[d] U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198-0001, USA
[e] Sigma Space Corporation, VIIRS Characterization Support Team (VCST), Lanham, MD 20706, USA

ARTICLE INFO

Article history:
Received 1 March 2014
Received in revised form 23 May 2014
Accepted 25 June 2014

Keywords:
Geostatistics
Nugget variance
Semivariogram
Signal-to-noise ratio
Spectral vegetation index
Standardized noise

ABSTRACT

In the past 40 years, many spectral vegetation indices have been developed to quantify vegetation biophysical parameters. An ideal vegetation index should contain the maximum level of signal related to specific biophysical characteristics and the minimum level of noise such as background soil influences and atmospheric effects. However, accurate quantification of signal and noise in a vegetation index remains a challenge, because it requires a large number of field measurements or laboratory experiments. In this study, we applied a geostatistical method to estimate signal-to-noise ratio (S/N) for spectral vegetation indices. Based on the sample semivariogram of vegetation index images, we used the standardized noise to quantify the noise component of vegetation indices. In a case study in the grasslands and shrublands of the western United States, we demonstrated the geostatistical method for evaluating S/N for a series of soil-adjusted vegetation indices derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. The soil-adjusted vegetation indices were found to have higher S/N values than the traditional normalized difference vegetation index (NDVI) and simple ratio (SR) in the sparsely vegetated areas. This study shows that the proposed geostatistical analysis can constitute an efficient technique for estimating signal and noise components in vegetation indices.

1. Introduction

A spectral vegetation index is a non-dimensional measure of spectral reflectances using an algebraic operation such as ratio, difference, weighted difference, or normalized difference in two or more bands to quantify vegetation biophysical characteristics. Visible and near-infrared (NIR) bands are the most commonly used wavelengths in development of spectral vegetation indices. A vegetation index based on the visible and NIR (VNIR) bands is generally related to photosynthetically active radiation (PAR) absorbed by vegetation canopies and is therefore considered a proxy for photosynthetic activity or vegetation greenness (Gamon et al., 1995; Myneni et al., 1995; Sellers, 1985).

Since the 1970s, various VNIR-based vegetation indices have been developed, which can be classified into five general groups: (1) simple vegetation indices using a ratio, a difference, or a normalized difference of red and NIR reflectances, for example, the simple ratio (SR) (Jordan, 1969), the difference vegetation index (DVI) (Tucker, 1979), and the normalized difference vegetation index (NDVI) (Rouse et al., 1974); (2) vegetation indices developed to adjust for the influence of background soil, for example, the perpendicular vegetation index (PVI) (Richardson and Wiegand, 1977), the weighted difference vegetation index (WDVI) (Richardson and Wiegand, 1977), the soil-adjusted vegetation index (SAVI) (Huete, 1988), the transformed soil-adjusted vegetation index (TSAVI) (Baret and Guyot, 1991), the modified soil-adjusted vegetation index (MSAVI) (Qi et al., 1994), the optimized soil-adjusted vegetation index (OSAVI) (Rondeaux et al., 1996), and the generalized soil-adjusted vegetation index (GESAVI) (Gilabert et al., 2002); (3) vegetation indices devised to compensate for atmospheric effects, such as the atmospherically resistant vegetation index (ARVI) (Kaufman and Tanré, 1992) and the global environmental monitoring index (GEMI) (Pinty and Verstraete, 1992); (4) vegetation indices combining corrections for both background soil influence and atmospheric effect, such as the soil atmospherically resistant vegetation index (SARVI) (Kaufman and Tanré,

* Corresponding author. Tel.: +1 6055946584; fax: +1 6055946529.
E-mail address: lji@usgs.gov (L. Ji).

http://dx.doi.org/10.1016/j.isprsjprs.2014.06.013
0924-2716/

1992), the modified normalized difference vegetation index (MNDVI) (Liu and Huete, 1995), the enhanced vegetation index (EVI) (Huete et al., 2002), and the two-band enhanced vegetation index (EVI2) (Jiang et al., 2008); (5) vegetation indices designed to increase the index's linearity with biophysical parameters, for example, the nonlinear vegetation index (NLI) (Goel and Qi, 1994), the renormalized difference vegetation index (RDVI) (Roujean and Breon, 1995), the modified simple ratio (MSR) (Chen, 1996), the green normalized difference vegetation index (GNDVI) (Gitelson et al., 1996), the green atmospherically resistant vegetation index (GARI) (Gitelson et al., 1996), the wide dynamic range vegetation index (WDRVI) (Gitelson, 2004), the linearized vegetation index (LVI) (Ünsalan and Boyer, 2004), and the linearized normalized difference vegetation index (LNDVI) (Jiang and Huete, 2010). In addition to the VNIR-based indices, indices using longer wavelengths (e.g., shortwave infrared) and narrow bands derived from hyperspectral sensors are also widely used.

An ideal vegetation index should meet two criteria. First, the index should be sensitive to the "signal" of a given biophysical parameter such as leaf area index, green vegetation fraction, PAR, or leaf chlorophyll concentration. Moreover, the index's sensitivity to signal should be consistent over the entire range of the biophysical parameter, requiring that the index not saturate for dense vegetation cover. Second, the index should be insensitive to "noise" such as the effects of background soil, atmosphere, canopy structure, sun-target-sensor geometry, and ground topography.

Evaluation and comparison of various vegetation indices have received great attention in the remote sensing community (e.g., Bannari et al., 1995; Gong et al., 2003; Xu et al., 2003; Silleos et al., 2006; Ji and Peters, 2007; Jiang and Huete, 2010). Some basic statistical techniques, such as correlation, regression, and analysis of variance, are useful in evaluating vegetation indices (e.g., Lawrence and Ripple, 1998; Purevdorj et al., 1998; Gong et al., 2003; Haboudane et al., 2004). To specifically estimate the signal and noise of a vegetation index, investigators devised several statistical metrics including the relative equivalent noise (Baret and Guyot, 1991), the vegetation equivalent noise (Huete et al., 1994), and the sensitivity function (Ji and Peters, 2007). These metrics, despite their different forms, can quantify the signal and noise components for a vegetation index based on the statistical relationship of the index to a biophysical variable. However, these methods for estimating signal and noise components require field-based measures of biophysical parameters, laboratory experiments, or model simulations.

Because a vegetation index consists of both desired signals and unwanted background noise, evaluation of a vegetation index can be simplified into an estimation of signal-to-noise ratio (S/N). In the remote sensing area, Smith and Curran (1999) summarized several image-based S/N evaluation methods: the homogeneous area method, the nearly homogeneous area method, the geostatistical method, the homogeneous block method, and the multiple waveband method. The geostatistical method was proposed by Curran and Dungan (1989) and later adapted or modified by Eklundh (1995), Atkinson et al. (1996), Atkinson (1997), Chappell et al. (2001), Foody et al. (2004), Atkinson et al. (2005), Atkinson et al. (2007), Guo and Dou (2008), Asmat et al. (2010), and others. The geostatistical method has been applied to estimate S/N of airborne and satellite images. For example, Curran and Dungan (1989) used this method to estimate S/N for the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data; Eklundh (1995), Chappell et al. (2001) performed S/N analysis for Advanced Very High Resolution Radiometer (AVHRR) NDVI data; Atkinson et al. (2005), Asmat et al. (2010) evaluated image noise for the hyperspectral images derived from Compact Airborne Spectrographic Imager; and Guo and Dou (2008) applied a modified geostatistical method to estimate S/N for the visible and infrared data acquired from China-launched FY-2 geostationary meteorological satellite series.

Although the geostatistical method has been used to estimate S/N of remotely sensed images, vegetation indices are different from regular images where each pixel records a brightness value that can be further converted to radiance or reflectance. The magnitude and sign of a vegetation index are normally irrelevant to the signal strength but are functions of land surface characteristics such as land cover types, vegetation density and condition, and soil properties. Therefore, not all the geostatistical metrics previously developed are suitable for vegetation indices. In this study, we applied the existing geostatistical techniques and proposed a procedure for estimating and comparing S/N values of different vegetation indices. In a case study in the grasslands and shrublands in the western United States we used several vegetation indices derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) data to demonstrate the effectiveness of our proposed S/N estimation method.

2. A brief review of geostatistical methods for S/N estimation

S/N in electronics is defined as a measure of signal strength relative to background noise. S/N can be expressed using different formulas because signal and noise can be defined in different ways (Schowengerdt, 1997). The most common definition of S/N in image processing is given by

$$S/N_{\mathrm{var}} = \frac{\sigma^2_{\mathrm{signal}}}{\sigma^2_{\mathrm{noise}}} \tag{1}$$

where $\sigma^2_{\mathrm{signal}}$ and $\sigma^2_{\mathrm{noise}}$ are the variances of signal and noise, respectively (Schowengerdt, 1997). The geostatistical estimation of S/N was developed based on the concept of the semivariogram, a key function in geostatistics. The sample semivariogram is defined as

$$\gamma(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} [z(x_i) - z(x_i + h)]^2 \tag{2}$$

where $z(x_i)$ is the value of a pixel location (x_i), h is the lag distance between pairs of pixels, m is the numbers of pairs of pixels at lag h, and $\gamma(h)$ is the estimate of the semivariogram at lag h (Isaaks and Srivastava, 1989). For semivariogram, there is an intrinsic stationarity assumption that the mean is a constant and the variance of the difference is the same everywhere in the region of interest. The sample semivariogram usually displays a characteristic shape, increasing from smaller to larger lags. The shape of a sample semivariogram is characterized by three parameters: sill variance (c), range (a), and nugget variance (c_0). In general, a semivariogram $\gamma(h)$ increases with large lags and levels off asymptotically. The semivariogram $\gamma(h)$ value and the lag h value at the asymptote are referred to as sill variance (c) and range (a), respectively. The nugget variance is the $\gamma(h)$ value when the lag h is zero. Ideally, $c_0 = 0$ for $h = 0$, but in reality, data noise or measurement error can cause a discontinuity at the origin of the semivariogram resulting in a positive nugget variance (Isaaks and Srivastava, 1989; Schowengerdt, 1997). A sample semivariogram can be fitted by a mathematical model, such as nugget effect model, spherical model, exponential model, and Gaussian model, to determine the semivariogram parameters c, a, and c_0. A complicated semivariogram model may contain two or more nested structures that combine multiple mathematical models (Isaaks and Srivastava, 1989).

The use of the nugget variance to estimate random noise for images was proposed by Curran and Dungan (1989). They justified this method by these two arguments: (1) the variance of an image is the sum of signal variance (or underlying variance) and the noise variance, and (2) when the lag approaches zero, the signal variance will be nearly zero, so the semivariogram of the image will consist of nearly pure noise variance. Curran and Dungan (1989) defined S/N_{mean} as