International Journal of Applied Earth Observation and Geoinformation 18 (2012) 515–527

Contents lists available at SciVerse ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Downscaling Landsat 7 ETM+ thermal imagery using land surface temperature and NDVI images

V. Rodriguez-Galiano[a,*], E. Pardo-Iguzquiza[b], M. Sanchez-Castillo[c], M. Chica-Olmo[a], M. Chica-Rivas[d]

[a] Departamento de Geodinámica, Universidad de Granada, 18071 Granada, Spain
[b] Instituto Geológico y Minero de España (IGME), Calle Ríos Rosas, 23, 28003, Madrid, Spain
[c] Departamento de Física Aplicada, Universidad de Granada, 18071 Granada, Spain
[d] Departamento de Análisis Matemático, Universidad de Granada, 18071 Granada, Spain

ARTICLE INFO

Article history:
Received 22 November 2010
Accepted 7 October 2011

Keywords:
Landsat
Land surface temperature (LST)
Thermal imagery
Downscaling Cokriging (DCK)
Image sharpening

ABSTRACT

Thermal infrared (TIR) satellite images and derived land surface temperature (LST) are variables of great interest in many remote sensing applications. However, the TIR band has a spatial resolution which is coarser than the other multispectral bands for a given satellite sensor (visible, near and shortwave infrared bands); therefore, the spatial resolution of the retrieved LST from available satellite-borne sensors is not accurate enough to be used in certain applications.

The application of a method is shown here for obtaining LST images with enhanced spatial resolution using the LST at a coarser resolution and the Normalized Difference Vegetation Index (NDVI) of the same scene using Downscaling Cokriging (DCK). A LST image with perfect coherence was obtained by applying this method to a Landsat 7 ETM+ image. This implies that, if the downscaled LST image is degraded to its original resolution, the degraded image obtained is identical to the original. Hence high spatial resolution LST images were obtained without altering the original radiometry with the inclusion of artefacts. Moreover, the performance of DCK was compared with global and local TSHARP methods. The RMSE of the sharpened images were 0.85, 0.92 and 1.1 K, respectively.

1. Introduction

It is well known that thermal infrared (TIR) sensors map the radiation emitted by the Earth in a range of 8–15 μm, while the solar reflective energy for that spectral range is, in general, negligible. The energy emitted by the Earth is related to the temperature and emissivity of the objects or materials on the ground. The TIR band registers a range of wavelengths with lower energy content than the other remote sensing spectral bands, implying that a TIR image usually has a coarser spatial resolution (see Table 1). Therefore, the spatial resolutions of these thermal images are coarser than those acquired at other wavelengths such as VIS, NIR, and SWIR. For instance, the Moderate Resolution Imaging Spectroradiometer (MODIS) provides thermal images at 1000 m resolution (near nadir), compared with the 250 m resolution for images acquired at other wavelengths. The ASTER TIR bands have a medium spatial resolution of 90 m, (15 m at the VIS/NIR bands), but there is a disadvantage to them as they are available only upon request. Hence, they cannot be used in studies demanding temporal continuity or which are of a retrospective nature. Therefore, if continuous medium-resolution and global coverage data are needed, only Landsat missions can offer these possibilities. For example, the TIR band (or band 6, 10.4–12.5 μm) of the Landsat Enhanced Thematic Mapper Plus (ETM+) satellite sensor has a spatial resolution of 60 m, whereas the other six multispectral bands (1, 2, 3, 4, 5 and 7, 0.45–0.52, 0.53–0.61, 0.63–0.69, 0.78–0.90, 1.55–1.75, 2.09–2.35 μm, respectively) have a spatial resolution of 30 m and the panchromatic band (or band 8, 0.52–0.90 μm) has a spatial resolution of 15 m (Goward et al., 2001).

Medium spatial thermal infrared remote sensing has been used to solve many practical environmental problems such as monitoring volcanic activity (Pieri and Abrams, 2005; Vaughan et al., 2010), detection of sea surface temperature and salinity (Zhang et al., 2004), climatological studies (Arnfield, 2003), evapotranspiration estimation (Allen et al., 2007; Anderson et al., 2008; Hong et al., 2009; Tang et al., 2009), the study of landscape ecological processes (Quattrochi and Luvall, 1999), earthquake predictions (Lü et al., 2000), monitoring of urban areas and urban processes (Gamba et al., 2005; Voogt and Oke, 2003), and land cover classification, among many other applications.

For many of the above said applications, the spatial resolution of currently available thermal imagery from the existing set of

* Corresponding author. Tel.: +34 958 243363; fax: +34 958 248527.
E-mail addresses: vrgaliano@ugr.es (V. Rodriguez-Galiano), e.pardo@igme.es (E. Pardo-Iguzquiza), mscastillo@ugr.es (M. Sanchez-Castillo), mchica@ugr.es (M. Chica-Olmo), mario.c.rivas@gmail.com (M. Chica-Rivas).

0303-2434/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jag.2011.10.002

Table 1
Technical characteristics of sensors with different spatial resolution between VNIR (visible and near infrared) and TIR (thermal infrared) bands.

Platform/Sensor	Spatial resolution VNIR bands (m)	Spatial resolution TIR bands (m)	Temporal resolution
Terra/MODIS	250–500	1000	1–2 days
Terra/ASTER	15	90	16 days
Landsat 5/TM	30	120	16 days
Landsat 7/ETM+	30	60	16 days

satellite-borne sensors is not sufficient, therefore, a need for image sharpening techniques to increase their resolution has emerged. Image sharpening refers to image processing techniques that combine two or more image sets from the same or different sensors, forming an enhanced final image (Wald, 1999; Wald et al., 1997; Yocky, 1996). A plethora of methods have been proposed for image sharpening: hue-intensity-saturation (HIS) method, principal components analysis (PCA) method, high-pass filter (HPF) method (Chavez et al., 1991), regression predictor (Price, 1999), smoothing filter based intensity modulation (Liu, 2000), wavelet transform and multiscale Kalman filter (Alparone et al., 2007; Gonzalez-Audicana et al., 2005; Nuñez et al., 1999; Otazu et al., 2005; Simone et al., 2002) and ARSIS method (Ranchin et al., 2003; Ranchin and Wald, 2000) to name just a few. More specifically, in the thermal domain, different methods have been used to enhance the spatial resolution of land surface temperature (LST) data derived from TIR images, using empirical relations with the same or other sensor higher resolution visible (VIS), near infrared (NIR), shortwave infrared (SWIR) and Normalized Difference Vegetation Index (NDVI) images (Agam et al., 2007; Gowda et al., 2007; Jeganathan et al., 2011; Kustas et al., 2003; Nichol and Wong, 2005; Stathopoulou and Cartalis, 2009; Tang et al., 2009; Yang et al., 2011; Zhan et al., 2011).

Thus, a logical aim of image processing is to devise methods for generating Landsat 7 ETM+ TIR images with a finer spatial resolution while preserving the spectral content of the image by using other high resolution images from the same or different sensor. The fused TIR image must provide a more accurate spatial description of variability in temperature of the ground objects but the downscaling process has to preserve the coherence in the radiometry of the original TIR band (Ranchin et al., 2003; Stathopoulou and Cartalis, 2009; Wald et al., 1997). The application of a highly consistent method, Downscaling Cokriging (DCK), was considered in this work. Although DCK has been proposed previously for pan-sharpening purposes by Pardo-Iguzquiza et al. (2006, 2011), prediction in the thermal domain is studied in this paper.

It has been previously acknowledged (Drury, 1987) that the correlation coefficient between the LST and the Landsat spectral bands may be considered low if compared with the correlation coefficient among the said spectral bands. However, it may be shown that there is a clear pattern of joint variability between the LST band and the other spectral bands as shown by statistical functions that measure joint variability in terms of the cross-covariance or cross-semivariogram. This function makes Cokriging an approximation with a high potential to incorporate only the spatial detail from the bands with the highest spatial resolution into the LST image. It can be justified by the statistical evidence of experimental cross-covariances and direct covariances or by the cross-semivariograms and direct semivariograms. Cokriging has additional advantages. Firstly, it takes pixel size explicitly into account (*support effect* in geostatistical jargon), as well as the point-spread-function of the sensor. The support effect (i.e. the change of pixel size) is referred to as upscaling or downscaling, depending on whether the pixel size becomes respectively increased or decreased. This paper focuses on downscaling, a method which implies that the downscaled image shows a higher variability – both in terms of space and radiometry – than the image with larger pixel sizes. Furthermore, this method is endowed with perfect coherence; the latter characteristic implying that if the high spatial resolution cokriged image is degraded by using the point-spread-function, the resulting image will be identical to the experimental image acquired by the satellite sensor. In the case of LST downscaling, this property is of paramount importance, as it has been found that over heterogeneous flat areas, LST at the coarse scale can be expressed as a simple areal average of the LST at the fine scale (Liu et al., 2006; Stathopoulou and Cartalis, 2009).

The purpose of this paper is to show the application and results of Cokriging in the context of merging NDVI and thermal images to obtain high resolution LST images from the Landsat 7 ETM+ sensor. In this manner, it is expected to contribute to increase the use of this variable (LST) in environmental applications which need high spatial resolution images. The performance of DCK has been evaluated as compared to that of the global TSHARP method (Agam et al., 2007; Kustas et al., 2003) and a variant of it, which takes different local neighbourhoods for the estimation into account (Jeganathan et al., 2011). It should be noted that the downscaling methodology is applicable to all the sensors presented in Table 1.

The general procedure to obtain the land surface temperature images with enhanced spatial resolution is sketched in Fig. 1.

1.1. Land surface temperature (LST) retrieval

The most suitable procedure to retrieve LST from a single band located in the TIR region is to invert the radiative transfer equation according to the following expression:

$$L_{\text{sensor},\lambda} = \left[\varepsilon_\lambda B_\lambda(T_s) + (1-\varepsilon_\lambda) L^{\downarrow}_{\text{atm},\lambda}\right] \tau_\lambda + L^{\uparrow}_{\text{atm},\lambda} \tag{1}$$

where L_{sensor} is the radiance measured by the sensor; ε is the land surface emissivity (LSE); $B(T_s)$ is the Planck radiation function; T_s is the land surface temperature (LST); $L^{\downarrow}_{\text{atm}}$ is the downwelling atmospheric radiance; τ is the total atmospheric transmissivity between the surface and the sensor; $L^{\uparrow}_{\text{atm}}$ is the upwelling atmospheric radiance and λ is the wavelength.

As inferred from Equation (1), in order to convert brightness temperature to surface temperature, the retrieval of emissivity is necessary. There are several methodologies for LSE computation (Sobrino and Raissouni, 2000; Valor and Caselles, 1996; Van De Griend and Owe, 1993). For Landsat series, land surface emissivity can be estimated from the NDVI threshold method (Sobrino et al., 2004, 2008). This method obtains the emissivity from the NDVI considering different cases:

$$\varepsilon_\lambda = \begin{cases} a_\lambda + b_\lambda \rho_{\text{red}}, & \text{NDVI} \prec \text{NDVI}_s \\ \varepsilon_v P_v + \varepsilon_s(1-P_V) + C_\lambda, & \text{NDVI}_s \le \text{NDVI} \le \text{NDVI}_v \\ \varepsilon_v + C_\lambda, & \text{NDVI} \succ \text{NDVI}_v \end{cases} \tag{2}$$

where ε_v and ε_s are the vegetation and soil emissivity, respectively, and P_v is the vegetation proportion obtained according to Carlson and Ripley (1997):

$$P_v = \left[\frac{(\text{NDVI} - \text{NDVI}_{\min})}{(\text{NDVI}_{\max} - \text{NDVI}_{\min})}\right]^2 \tag{3}$$

Inversion also requires knowledge of the atmospheric parameters τ, $L^{\downarrow}_{\text{atm}}$ and $L^{\uparrow}_{\text{atm}}$, which is not always possible. In order to solve this problem, the said parameters can be computed through a radiative transfer code such as MODTRAN (Berk et al., 1998). Additionally, variables such as air temperature, pressure, water vapour content, etc., obtained from *in situ* radio soundings or modelled atmospheric profiles, can be incorporated for the computation of these parameters.