

This is an author produced version of a paper published in *Remote Sensing of Environment*. This paper has been peer-reviewed and is proof-corrected, but does not include the journal pagination.

Citation for the published paper:

Reese, H., Olsson, H. (2011) C-correction of optical satellite data over alpine vegetation areas. *Remote Sensing of Environment*. Volume: 115
Number: 6, pp 1387-1400.

<http://dx.doi.org/10.1016/j.rse.2011.01.019>

Access to the published version may require journal subscription.
Published with permission from: Elsevier



Epsilon Open Archive <http://epsilon.slu.se>

C-correction of optical satellite data over alpine vegetation areas: A comparison of sampling strategies for determining the empirical c -parameter

Heather Reese and Håkan Olsson

Section of Forest Remote Sensing
Department of Forest Resource Management
Swedish University of Agricultural Sciences
Umeå, Sweden

Abstract

Semi-empirical topographic normalization methods (e.g., C-correction) have been widely used to correct illumination differences in optical satellite data. The objective of this study was to examine the precision and accuracy of the C-correction's empirical parameter, c , as a function of the sample from which it was derived. Three sampling methods were compared: a random sample, a sample stratified on north and south aspects, and a sample stratified by cosine of the solar incidence angle, i . In the latter, power allocation was used to determine the quantity of observations for each stratum. Four overlapping satellite images were used (two Landsat 5 TM and two SPOT 5 HRG) with different acquisition dates and large solar zenith angles over an alpine region in Sweden. The sample stratified by cosine of i produced c with the highest precision from repeated trials and had coefficients of determination (R^2) twice as high as those from the other sampling methods. Use of power allocation in the cosine of i stratified sample enabled better representation of spectral variability; this was particularly important for the NIR band where the outcome of c differed according to sampling method. Evaluations using t -tests and classification accuracy showed that c derived from the cosine of i stratified sample correctly normalized a larger percentage of the evaluation data. The distribution of cosine of i in the study area, the spectral variability and vegetation types exert influences to consider when sampling to derive c . Although sampling was restricted to alpine vegetation only, some

vegetation classes may have benefitted from separate c -parameter calculation. In general, dry alpine heath and alpine grass heath had relatively higher c -parameters, mesic alpine heath was slightly lower, and alpine willow and alpine meadow had lower c -parameters for the near-infrared band. The cosine of i stratified sampling method using power allocation may be useful for calculation of c for vegetation conditions other than those presented here, as well as for other empirical parameters (e.g., Minnaert k).

1 Introduction

1.1 Background

The radiance recorded by a satellite sensor is affected by several factors, including topography, solar illumination angles, atmosphere, and vegetation characteristics. Topographic characteristics of the landscape, such as slope and aspect, in combination with the solar zenith and azimuth angles, result in illumination differences apparent within a satellite image. Topographic normalization methods adjust the spectral radiance in an image so that a vegetation class will have similar spectral values whether facing away from or towards the sun (Holben and Justice 1980). There are several categories of topographic normalization methods, such as photometric/photometric-empirical, statistical-empirical, sun-canopy-sensor, and physically based models (Soenen et al. 2008). The photometric-empirical category includes the “semi-empirical” methods, which consist of a photometric function modified by an empirical parameter. Examples of semi-empirical methods are the Minnaert correction (Smith et al. 1980) and C-correction (Teillet et al. 1982), which use the empirical parameters k and c , respectively. These empirical parameters are derived differently, yet both are based on the relationship between the image’s spectral data and the topographic and illumination characteristics, and are determined from a sample of these data.

Despite numerous topographic normalization studies, details or guidelines about the sample used to calculate the empirical parameters have seldom been provided. The aim of this study was to examine the resulting empirical parameter as a function of the sample from which it was derived. In the literature, one of the few recommendations regarding an empirical parameter’s sample can be found in Civco (1989) who stratified a sample by aspect with “large samples of an equal number of pixels ($n = 1,390$) falling on northern and southern slopes”. In general, two categories of sampling methods for

calculating semi-empirical parameters can be distinguished in the literature: 1) selection (often subjective) of a relatively small number of observations (generally $n < 100$) for a target vegetation type over a range of topographic conditions (e.g., Blesius and Weirich 2005; Ekstrand 1996; Smith et al. 1980), or 2) random sampling of a varying quantity of observations from either a subset or an entire image. For example, Teillet et al. (1982) sampled $n = 1,265$ and $n = 1,038$ in two Landsat images to calculate c over five forest types; Colby (1991) sampled all pixels and a subset from a forested Landsat Thematic Mapper (TM) image to calculate global and local k ; and, Bishop and Colby (2002) compared k calculated globally, locally and specifically for three broad land-cover types using SPOT 3 images. The work described in the present study focused on use of random sampling, the second of the two methods mentioned, to collect the sample used to determine c .

Many topographic normalization studies have the objective of comparing methods or modifying them to improve their performance. While different semi-empirical methods have been found to have similar results (Meyer et al. 1993; Richter et al. 2009), there have been general problems such as over-correction of extreme slopes (Riaño et al., 2003), low correlation between spectral data and illumination angles (Bishop and Colby 2002; Carpenter et al. 1999), and unsatisfactory corrections. Among the modifications are slope-specific corrections (Ekstrand 1996; Nichol et al. 2006), slope-smoothing (Kobayashi and Sanga-Ngoie 2009), and use of different models for the visible versus infrared bands (Richter et al. 2009; Vincini and Frazzi 2003). Some studies have touched upon problems with the empirical parameter, such as Ekstrand (1996) who found an increased k was needed for better correction of the near-infrared (NIR) band, Civco (1989) who found difficulties in normalizing the NIR band in particular, and Gu and Gillespie (1998) who found c -parameters to be so large in general that they “lack an exact physical explanation”. One development of methods has been to apply separate topographic normalizations specific to the vegetation type. This is based on the idea that non-Lambertian reflectance varies in degree with surface roughness and therefore by vegetation characteristics (Holben and Justice 1980; Teillet et al. 1982). Gu and Gillespie (1998) suggested separate corrections for forest and non-forest, using a sun-canopy-sensor model for forest and C -correction for non-forest. Bishop and Colby (2002) derived individual empirical parameters for three broad land cover classes, namely vegetation, non-vegetation and snow. Other researchers have derived empirical parameters for specific classes, such as coniferous forest species (Teillet et al. 1982) or separate normalizations for forests with differing canopy complexity (Kane et al. 2008). Bishop and

Colby (2002) and Soenen et al. (2008) concluded that more study was needed regarding vegetation-specific topographic normalization.

The present study focused on alpine vegetation, which has different reflectance characteristics than forests, for which the majority of topographic normalization studies have been conducted. The geotropic growth of trees and resulting canopy self-shadow within forests influences the observed radiance, especially on sloping terrain (Gu and Gillespie 1998). Although alpine vegetation is shorter in stature, the reflectance behavior can also be complex and non-Lambertian (Hugli and Frei 1983; Vierling et al. 1997). In general, the radiance within a pixel can be influenced by the mixture of background reflectance (e.g., soil), non-photosynthetic vegetation, green vegetation, and shadow (Roberts et al. 1993). Studies of alpine vegetation have found that the amount of green live foliar phytomass has a strong influence on NDVI (Riedel et al. 2005; Stow et al. 1993). Additionally, there is a strong correlation between alpine vegetation classes, topographic conditions and NIR (and NDVI) reflectance (Deng et al. 2007). The Swedish alpine area is covered by a mosaic of vegetation types (e.g., willow, dwarf shrubs, grasses, forbs) varying in density, shape and height (from < 1 m to 2 m, by definition). Given the vegetation dependence of topographic normalization, the effect of this heterogeneous mixture of alpine vegetation on the derivation of an appropriate empirical parameter for topographic normalization must be considered.

The objective of this study was to examine the precision and accuracy of the empirical parameter as a function of the sample from which it was derived. The sampling method and the quantity of sample observations (determined using sampling theory) were hypothesized to be of importance to the resulting empirical parameter. Three sampling methods for calculating the empirical parameter were tested: a random sample; a stratified random sample with stratification on north and south aspects; and, a stratified random sample with stratification by the cosine of the solar incidence angle, i . For the sample stratified by cosine of i , an optimal allocation method called power allocation (Bankier 1988) was used to determine the quantity of observations for each stratum. The work was carried out in an alpine area at a high northern latitude in Sweden (66° 00' N, 15° 30' E) where solar zenith angles may be large. Four overlapping satellite images (two Landsat 5 TM and two SPOT 5 HRG) were used, each with different acquisition dates and varying illumination conditions. In this study, we have concentrated on use of the semi-empirical C-correction topographic normalization method, a description of which follows.

1.2 The C-correction

The C-correction consists of a modified cosine correction plus the empirical parameter c (Teillet et al. 1982), which is derived from the linear relationship between the spectral data and the cosine of the solar incidence angle, i , with respect to surface normal. Cosine of i ($\cos i$) is calculated as a function of the local terrain slope and aspect, and the solar illumination angles upon the surface at the time of satellite data acquisition (Table 1, Eq. 1). Linear regression is used to estimate intercept (b) and gradient (m), using $\cos i$ as the independent variable and reflectance ($\hat{\rho}_{\lambda t}$) as the dependent variable (Eq. 2). The c -parameter is calculated as b divided by m (Eq. 3) for each wavelength band since the relationship between reflectance and $\cos i$ is wavelength dependent (Teillet et al. 1982). The c -parameter is added to the numerator and denominator of the cosine correction to form the C-correction equation (Eq. 4). Teillet et al. (1982) state that c emulates the effect of path radiance, but that the analogy is not exact. The angles described in Table 1 are depicted in Figure 1.

Table 1

Symbols and their definitions (a) used in the equations (b) for C-correction (Teillet et al. 1982).

a) Symbols and definitions	
i	= solar incidence angle with respect to surface normal
z	= solar zenith angle
s	= terrain slope angle
a	= solar azimuth angle
a'	= terrain aspect angle
$\hat{\rho}_{\lambda t}$	= topographically influenced (t) reflectance of band λ
b	= intercept of linear regression
m	= gradient of linear regression
c_{λ}	= c -parameter for band λ
$\hat{\rho}_{\lambda h}$	= topographically normalized (h) reflectance of band λ
b) Equations	
Cosine of i ($\cos i$)	$\cos i = \cos z * \cos s + \sin z * \sin s * \cos (a - a')$ (1)
Linear regression between reflectance and $\cos i$	$\hat{\rho}_{\lambda t} = b + m * \cos i$ (2)
c -parameter	$c_{\lambda} = \frac{b}{m}$ (3)
C-correction	$\hat{\rho}_{\lambda h} = \hat{\rho}_{\lambda t} \frac{\cos z + c_{\lambda}}{\cos i + c_{\lambda}}$ (4)

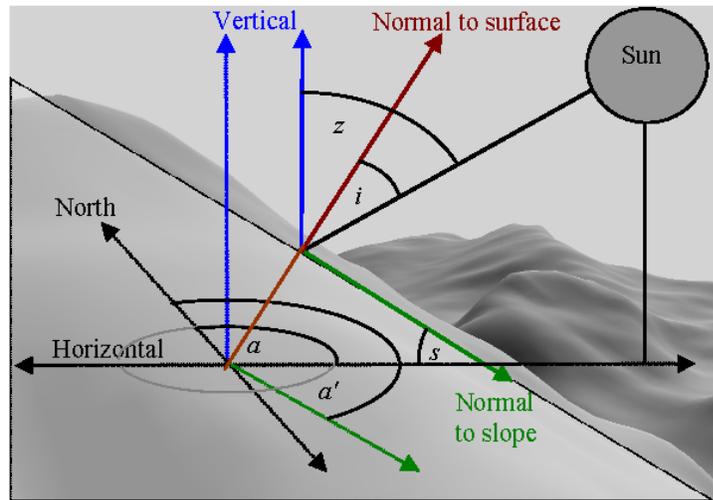


Fig. 1. Illustration of topographic and solar illumination angles, including solar azimuth (a), terrain aspect (a'), solar zenith (z), terrain slope (s), and solar incidence with respect to surface normal (i).

2 Materials

2.1 Study Area

The 110 x 110 km study area was in a mountainous region of northern Sweden (Fig. 2) with elevations ranging from 295 to 1768 m. The geology of the area is highly varied bedrock, including schist, gneiss, phyllite, and amphibolites. The mean annual temperature here is -1°C, the mean July temperature of +11°C, and there is between 800 to 1200 mm precipitation annually. These factors and others give rise to a varied alpine vegetation, consisting of a mosaic of several species of dwarf shrubs, willow, other shrubs, low-growing heath vegetation, forbs, grasses, and numerous mosses and lichens. The alpine vegetation classes used in this study are given in Table 2. The class definitions follow Rafstedt (1983), classified by vegetation height and composition, and with occurrence strongly related to topographic influences such as altitude, slope and moisture regime. Sub-alpine birch (*Betula pubescens ssp. czerepanovii*) trees are defined as being 2 to 5 m in height, and form the tree limit at around 600 to 800 m elevation, while coniferous forest (primarily *Picea abies*) generally reaches up to 500 to 600 m elevation.

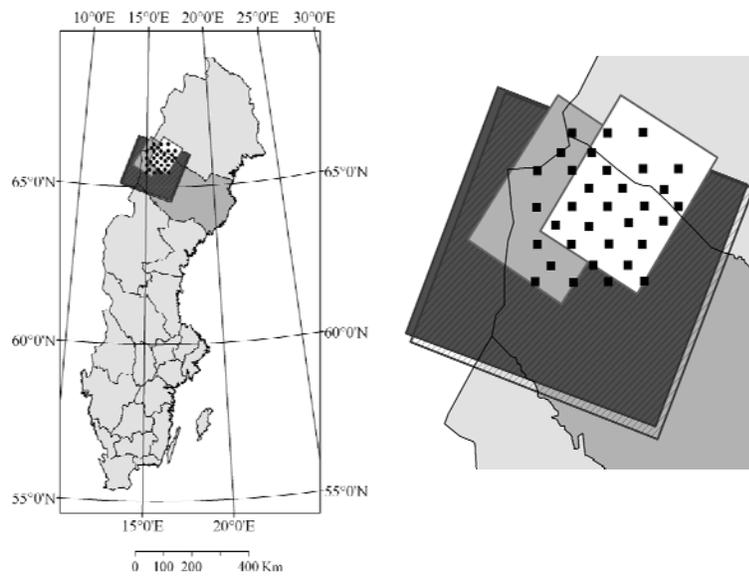


Fig. 2. Study area in Sweden (left) with Västerbotten province in darker grey. Enlargement (right) showing the two overlapping Landsat TM images from 2006 and 2005, the SPOT 2008 two-image mosaic (gray), the SPOT 2004 two-image mosaic (white), and aerial photography sampling units (small black squares).

Table 2 Classes, definitions and vegetation height of the alpine vegetation classes in the evaluation dataset. The number of plots in the evaluation dataset for each class is given, as well as the mean elevation above sea level. The class assigned to the plot has 70% or greater area coverage within the plot (10m radius). Plots with < 70% area coverage of a single class were not included in the evaluation dataset.

Class name	Definition	Vegetation height (dm)	Elev. mean (m)	Nr. Plots
Grass heath	Low-growing grasses on poor soil, where > 10% of plot area is grass or half-grass. Often three-leaved rush (<i>Juncus trifidus</i>), stiff sedge (<i>Carex bigelowii</i>), sheep's-fescue (<i>Festuca ovina</i>), alpine hair-grass (<i>Deschampsia alpina.</i>), sweet vernal-grass (<i>Anthoxanthum odoratum</i>), mat-grass (<i>Nardus stricta</i>), etc. Most often in middle alpine region.	< 4	1003	114
Extremelydry heath	Extreme exposure to wind, often bare of snow during winter, with sparse, low-growing vegetation. Often crowberry (<i>Empetrum hermaphroditum</i>), mountain bearberry (<i>Arctostaphylos alpinus</i>), trailing azalea (<i>Loiseleuria procumbens</i>), diaspensia (<i>Diapensia lapponica</i>), and lichens. In middle to low alpine region.	< 1	926	28
Dry heath	Low-growing shrubs dominate over grasses or forbs. Often crowberry (<i>Empetrum hermaphroditum</i>) and dwarf birch (<i>Betula nana</i>). In low alpine region.	< 3	878	118
Short alpine meadow	Forbs and grasses dominate over shrubs. Often meadow buttercup (<i>Ranunculus acris</i>), alpine lady's mantle (<i>Alchemilla alpina</i>), mountain violet (<i>Viola biflora</i>), sweet vernal-grass (<i>Anthoxanthum odoratum</i>), alpine's cat-tail (<i>Phleum alpinum</i>), etc. In low alpine region.	< 5	857	118
Alpine willow	Different species of willow (<i>Salix lapponum</i> , <i>Salix glauca</i> , <i>Salix lanata</i> , and other <i>Salix spp.</i>) with at least 25% willow coverage of plot area. In low alpine region.	5– 20	802	143
Mesic heath	Shrubs dominate over grasses or forbs. Often blueberry (<i>Vaccinium myrtillus</i>), dwarf birch (<i>Betula nana</i>), blue heath (<i>Phyllodoce caerulea</i>), willow (<i>Salix spp.</i>), juniper (<i>Juniperus communis</i>). In low alpine region.	3– 10	799	158
Tall alpine meadow	Forbs and grasses dominate over shrubs. Globeflower (<i>Trollius europaeus</i>), alpine blue-sow-thistle (<i>Cicerbita alpina</i>), cow-parsley (<i>Anthriscus sylvestris</i>), monk's-hood (<i>Aconitum napellus</i>), ferns, etc. In low alpine region.	5– 20	770	14
				693

2.2 Satellite images

Four sets of satellite data acquisitions were used in this study: two Landsat 5 TM images from 2006 and 2005, and two SPOT 5 High Resolution Geometric (HRG) XS image mosaics from 2008 and 2004 (Table 3). The Landsat 5 TM data had a spatial resolution of 25 m and TM bands 1-5 and 7 were used. The SPOT 5 HRG data had a spatial resolution of 10 m and XS bands 1-4 were used. Each SPOT 5 mosaic consisted of two contiguous images acquired in the same satellite pass by the HRG-1 sensor. By using images acquired with different solar zenith and azimuth angles we aimed to test the calculation of c for varying illumination conditions. In this study, the SPOT 5 images had large positive view angles (i.e., sensor viewing eastward). The Landsat 2006 image was cloud free, while the Landsat 2005 and SPOT 2008 images had some clouds (<10%), and the SPOT 2004 image had jet contrails on the western edge. The images were registered to the Swedish RT90 coordinate system, with less than one-half pixel root mean square error (RMSE) and good co-registration between all images. All images were from the summer vegetation period but from different years, with potential differences in phenology or soil moisture conditions, although such influences were not noticeable in the images used. Terra/Aqua MODIS Nadir-/BRDF-adjusted reflectance 16-day composites corresponding to the same date of the Landsat or SPOT images were acquired for the purpose of reflectance normalization.

Table 3
Characteristics of the satellite images used in the study.

Sensor	Acquisition Date	Path/Row	Spatial Resolution	Local Time GMT+1	Solar azimuth angle	Solar zenith angle	Sensor viewing angle
Landsat 5 TM	2006 Aug 19	197/014	25 m	11.16	167°	53°	Nadir
Landsat 5 TM	2005 July 31	197/014	25 m	10.57	163°	49°	Nadir
SPOT 5 HRG-1	2008 Aug 24	042/214-0 & 042/213-0	10 m	12.19	186°	55°	+26.8°
SPOT 5 HRG-1	2004 July 29	046/213-0 & 046/213-8	10 m	12.27	189°	47°	+26.7°

2.3 Digital Map Data

Digital map data used included the Swedish National Mapping Agency's Digital Elevation Model (DEM), the "Road Map," and the "GSD-Land and Vegetation Cover" map. The 50m resolution DEM, made from manual stereo interpretation of aerial photos, was the best full coverage DEM available for the study area. The Road Map contains land cover data at 1:100,000 scale with broad classes such as alpine vegetation, sub-alpine forest, non-alpine forest, open land (non-alpine), urban, water, wetland, and glaciers. The Road Map's "alpine vegetation" class could, in reality, contain bare rock, dry alpine heath, mesic alpine heath, alpine grass heath, alpine meadow, alpine willow, snow, and snow bed vegetation. The GSD-Land and Vegetation Cover map is derived from a classification of Landsat data and has more detailed vegetation classes than the Road Map. In the alpine region, the thematic classes are grass heath, other heath (i.e., dry/mesic heath), thickets (i.e., willow), alpine meadow, and sparse vegetation (i.e., < 50% vegetation cover).

2.4 Evaluation dataset

The evaluation dataset was based on a nationwide, systematic two-stage sample from the "National Inventory of Landscapes in Sweden" (NILS) program (Ståhl et al. 2010). NILS' primary sampling units are 5 x 5 km squares with 25 km distance between them in the mountain areas. The primary sampling unit is covered by three color-infrared aerial photographs at 1:30 000 scale with 60% overlap using a Zeiss-Intergraph Digital Mapping Camera (Z/I DMC). We wanted to double the number of primary sampling units available for this study, and therefore acquired supplemental aerial photographs with 12.5 km distance diagonal to the NILS primary units (Fig. 2). This resulted in 32 primary sampling units total in the Landsat images, and 19 and 22, respectively in the SPOT 2004 and 2008 image mosaics. The secondary sampling unit was a 10 m radius photo-interpretation plot, of which there were 110 plots systematically arranged (ten rows with 500 m between rows) within each 5 x 5 km primary sampling unit. The secondary sample plots were located at the center of the satellite image pixel, for both SPOT and Landsat. Vegetation classes and percent cover per class were determined for the center point and the entire 10 m radius plot via stereo interpretation using a digital photogrammetric workstation; a subset of plots were field-checked. In this study we concentrated on the seven classes listed in Table 2, omitting spectrally different classes such as snow/ice, water, wet heath, wetland, snow bed vegetation, and bare rock. Only plots with 70% or greater coverage of a single vegetation class within the plot were used with the intent of excluding

pixels with a highly mixed spectral signature. Upon reduction, the total number of secondary sampling units in the evaluation dataset was 693. The evaluation data were based on a systematic sample, and therefore provided a representation of class frequency that may differ from other sampling methods, for example, from an evaluation dataset stratified on a topographic parameter (e.g., elevation or *cosi*).

3 Methods

3.1 Data pre-processing

3.1.1 Reflectance normalization of the image data

Performing topographic normalization on atmospherically corrected reflectance data is recommended (Shepherd and Dymond 2003; Teillet 1986). In this study, the Landsat and SPOT data were reflectance-normalized relative to Terra/Aqua MODIS Nadir-/BRDF-adjusted reflectance data from a corresponding acquisition date (Olthof et al. 2005; Reese et al. 2009; Roy et al. 2008). The choice of sampling plots and method for reflectance normalization is important (Vicente-Serrano et al. 2008); we used linear regression with randomly sampled plots ($n = 5,000$). Pixels with water, wetlands, agriculture, urban areas, clouds, and snow/ice were masked and excluded from the reflectance normalization process. The Landsat and SPOT images were subset to match the MODIS pixel corner coordinates and block aggregated to a 500 m pixel. Band-wise linear regression of the Landsat or SPOT digital number (DN) and the corresponding MODIS reflectance was done according to Eq. 5,

$$\rho_{\lambda, modis} = \beta_0 + \beta_1 * DN_{\lambda} \quad (5)$$

where $\rho_{\lambda, modis}$ was the reflectance value for band λ of MODIS, and DN_{λ} was the corresponding digital number for band λ of Landsat or SPOT. The resulting regression equation provided an intercept (β_0) and gradient (β_1), used to reflectance normalize the Landsat and SPOT images with Eq. 6,

$$\hat{\rho}_{\lambda t} = \beta_0 + \beta_l * DN_{\lambda} \quad (6)$$

where $\hat{\rho}_{\lambda t}$ was the normalized reflectance for band λ of the Landsat or SPOT image (yet still influenced by topography, as indicated by subscript t).

3.1.2 Stratifying the image to create the alpine vegetation subset

The full Landsat and SPOT images contained several different classes (e.g., alpine vegetation, sub-alpine forest, non-alpine forest, water, wetland, urban, and agriculture) from which we focused on the alpine vegetation. Following the premise that topographic normalization is vegetation type dependent, the 1:100 000 “Road Map” was used to mask each satellite image into four separate image subsets, namely alpine vegetation, sub-alpine forest, non-alpine forest, and other (water, wetland, urban, agriculture). All topographic normalization work described in this paper was carried out upon data from the alpine vegetation subset only. Bare rock and snow/ice occurred within the alpine vegetation subset, but have reflectance properties differing from vegetated surfaces (Giardino and Brivio 2003; Hugli and Frei 1983) and were therefore masked away from the alpine vegetation subset using a threshold of $NDVI < 0.3$.

3.1.3 DEM-derived data

Slope, aspect and $cosi$ (Table 1, Eq. 1) were derived from the 50 m DEM. $Cosi$ values generally range from -1 to 1, where negative values are completely shadowed slopes (e.g., vertical overhangs with no direct solar illumination), while positive values range from lower values sloping away from the sun to higher values sloping towards the sun. The $cosi$ raster was interpolated from 50 m into 10 m and 25 m resolution rasters (Fig. 3) using natural neighbor interpolation (Bater and Coops 2009; Sibson 1981). Comparison between the interpolated DEMs and field-measured GPS elevations ($n = 360$) showed that the original 50 m DEM had an RMSE of 9.0 m for elevation, while the 10 m interpolated DEM had an RMSE of 7.9 m for elevation. Within the alpine vegetation subset, $cosi$ followed a normal distribution, namely $N(0.583, 0.013)$, $N(0.633, 0.011)$, $N(0.562, 0.014)$, and $N(0.660, 0.010)$, for the Landsat 2006 and 2005 images and the SPOT 2008 and 2004 images, respectively.

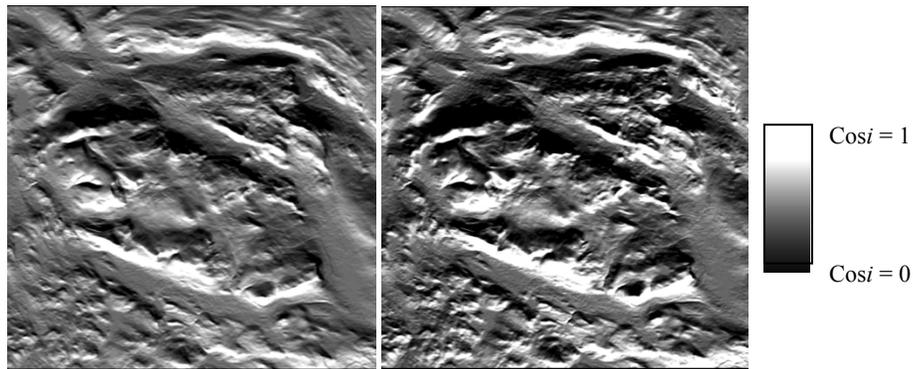


Fig. 3. Cosine of i for a subset of SPOT 2008 to left (solar zenith = 55° , solar azimuth = 186°) and Landsat TM 2006 to right (solar zenith = 53° , solar azimuth = 167°). For reference, satellite image subsets of the area shown here appear in Fig 6.

3.2 Topographic normalization

3.2.1 C-correction

In this study, we topographically normalized the satellite data using the C-correction, as described in Section 1.2. Additional modifications to compensate for radiation reflected from adjacent areas (Proy et al. 1989) or cast shadows (Giles 2001) were not incorporated. Determination of c was tested by using different sampling methods, a description of which follows.

3.2.2 Sampling methods to determine the c-parameter

We tested three strategies to sample the data for calculation of c :

1. a random sample with a large number of observations, called the “random sample”;
2. a stratified random sample where observations were stratified by aspect (north- and south-facing slopes), called the “aspect sample”; and,
3. a stratified random sample where observations were stratified by cosine of i , called the “ cosi sample.”

All samples were drawn from the alpine vegetation image subset only. A broad definition of north (315° - 360° and 0° - 45°) and south aspects (135° - 225°) was used for the aspect sample. For the cosi sample, the stratification was based on 0.1 increments of positive cosi values, resulting in ten strata (e.g., 0.01-0.10, 0.11-0.20...0.91-1.00).

3.2.3 Quantity of observations in the sample

In topographic normalization studies, there is often mention of taking a “large” random sample with which to calculate the empirical parameter. The term “large,” however, is relative and can vary in quantity according to the particular study. Sampling theory can provide guidance regarding the number of observations needed to make an estimate with a designated precision. Thigpen (1987) suggested that when a sample is taken from a bi-variate normal distribution for the purpose of linear regression, three parameters can help determine sample size (n): 1) the correlation (r) between the two variables, 2) a specified acceptable relative error (δ) for estimating the slope of the regression, and 3) the Student’s t -statistic for a specified α -level. If the correlation (r) is known or can be determined by a sample of the data, and desired error (δ) and probability levels specified, n can be solved for using Eq. 7 (Thigpen, 1987).

$$(n-1)^{0.5} = \frac{t_{\alpha/2} \left[\frac{(1-r^2)}{r^2} \right]^{0.5}}{\delta} \quad (7)$$

When correlation between variables is relatively low (e.g., $r = 0.3$), with $\delta = 0.05$ and $\alpha = 0.05$, Eq. 7 indicates that $n = 16,200$ is necessary, while for $r = 0.5$, $n = 4,600$. In the present study, a sample of the spectral data within the alpine vegetation subset indicated that the correlation between cosi and the visible bands was around $r = 0.3$, but higher for the NIR and SWIR bands ($r \geq 0.5$). Therefore, for the “large” random sample, we chose an $n = 16,500$ sample for all bands. To study the effect of reducing the number of plots on calculating c , we also sampled with $n = 5,000$, and $n = 1,600$ since the sample size referred to in the literature as being “large” has often been $n < 2,000$. The total number of observations used for the aspect and cosi stratified samples was smaller ($n = 5,000$) than that for the largest random sample ($n = 16,500$), as the correlations were higher.

3.2.4 Observation allocation (including power allocation) in the stratified random samples

For the sample stratified by aspect, the random observations within the north and south aspects had equal allocation ($n = 2,500$ for each stratum). For the cosi sample, we used a “power allocation” (Bankier 1988; Lehtonen and Pahkinen 2004), which is a modified form of optimal allocation, to determine the number of observations per stratum. Power allocation allows small but important strata to be given weight, and incorporates descriptive statistics (e.g., standard deviation or coefficient of variation) about the variable of interest.

The power allocation used here was based on Bankier's (1988) power allocation (Eq. 8), for a dataset with a total of $h = 1 \dots L$ strata, and a given total sample size, n .

$$n_h = n \times \frac{N_h^q \times CV_{hy}}{\sum_{h=1}^L N_h^q \times CV_{hy}} \quad (8)$$

The power allocation in Eq. 8 provides a solution for a minimum sample size (n_h) for stratum h , where N_h is the stratum population size (i.e., total number of pixels in stratum h), and CV_{hy} is the coefficient of variation ($CV = \sigma / \mu$) of the variable of interest, y (e.g., the reflectance), within stratum h . The coefficient q refers to the desired power, where $1 \geq q \geq 0$, and is assigned by the user. Solutions of $q = 0.3$ or 0.5 are often used (Lehtonen and Pahkinen 2004). When $q = 1$, the result is similar to a Neyman allocation, which is a function of stratum size (Lehtonen and Pahkinen 2004; Neyman 1934), while if $q = 0$, stratum size is effectively not considered. A value of q between 0 and 1 is a compromise between these allocations (Bankier 1988). In this study, we used $q = 0.3$ for all *cosi* strata for the visible bands, while for the NIR and SWIR bands, we have suggested a different assignment of q . In particular, this appeared to be important for the NIR band, which exhibited a large standard deviation and CV for the spectral values in the higher *cosi* strata, and had potentially very low spectral values occurring within the low *cosi* stratum. For this reason, a larger q ($q = 0.4$) was used for the NIR and SWIR bands in the highest *cosi* stratum ($cosi \geq 0.9$), and a lower q ($q = 0.2$) in the lowest *cosi* stratum ($0 < cosi \leq 0.1$). Strata adjacent to these had a 0.05 increment to reach to $q = 0.3$, which was used for all other strata (as shown by example in Table 4).

Table 4

An example from the Landsat TM 2006 image (alpine vegetation subset only). Mean, standard deviation and total number pixels for the red, NIR and SWIR1 bands for each cosine of i stratum. For the $cosi$ sample ($n = 5,000$) with power allocation (Eq. 8), the table shows q as used in power allocation for visible and IR bands, and the resulting sample size per $cosi$ stratum (n_h) for red, NIR and SWIR1.

Cosi stratum	μ Red Reflect	μ NIR Reflect	μ SWIR Reflect	σ Red Reflect	σ NIR Reflect	σ SWIR Reflect	Total pixels (N_h)	q Vis	q IR	n_h Red	n_h NIR	n_h SWIR1
.9-1.0	0.088	0.322	0.326	0.024	0.075	0.042	49862	0.3	0.4	252	845	517
.8-9	0.077	0.323	0.303	0.022	0.069	0.041	279664	0.3	0.35	433	718	516
.7-8	0.068	0.309	0.276	0.018	0.056	0.039	1218092	0.3	0.3	637	581	510
.6-7	0.060	0.282	0.242	0.015	0.044	0.038	4333359	0.3	0.3	866	714	814
.5-6	0.055	0.262	0.218	0.013	0.037	0.033	4455306	0.3	0.3	829	671	791
.4-5	0.052	0.243	0.199	0.013	0.036	0.025	1503108	0.3	0.3	632	502	477
.3-4	0.045	0.223	0.170	0.011	0.032	0.030	354177	0.3	0.3	410	319	437
.2-3	0.036	0.195	0.135	0.011	0.041	0.036	91097	0.3	0.3	324	305	444
.1-2	0.026	0.140	0.086	0.010	0.047	0.037	34812	0.3	0.25	336	239	350
<.1	0.019	0.104	0.056	0.007	0.026	0.019	26986	0.3	0.2	276	101	139
Total	0.058	0.271	0.230	0.016	0.049	0.046	12346493	--	--	5000	5000	5000

3.2.5 Thresholding complete shadow from the sample

There may be observations in the sample where extremely low reflectance values occur in low $\cos i$ areas (e.g., $0 < \cos i \leq 0.2$), potentially resulting in a sample with a non-linear relationship (e.g., Fig. 4). These may be pixels with no reflectance, located in complete shadow, and erroneously labeled with positive $\cos i$ values. This may occur in images with large solar zenith angles, next to steep anti-solar-facing slopes with negative $\cos i$ values, or may be due to the limited resolution or accuracy of the DEM (Conese et al. 1993). To remove these from the sample, observations with true negative $\cos i$ were used to determine a minimum threshold for identifying completely shadowed observations (Fig. 4). The NIR band was used to determine the threshold value, however, observations were thresholded from the samples for all wavelength bands. These observations were checked manually to confirm that they represented complete shadow, and the threshold was applied conservatively so as not to eliminate valid observations within low $\cos i$ strata.

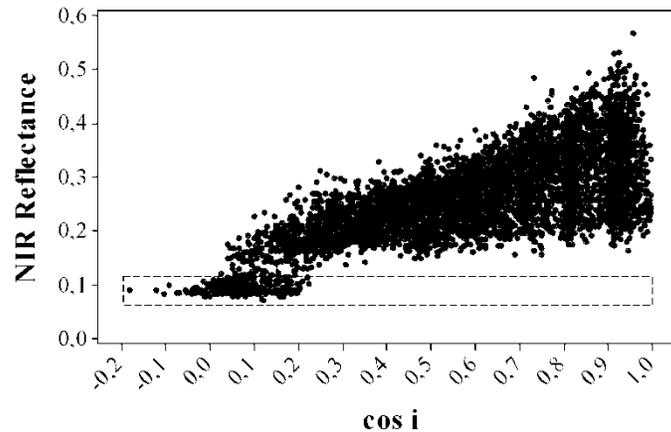


Fig. 4. $\cos i$ sample from Landsat TM 2006 image, containing observations in complete shadow ($\cos i < 0$). The dashed box indicates observations below the NIR threshold to be removed.

3.3 Assessment

3.3.1 Assessing precision of the c -parameter

To assess the repeatability (i.e., precision) of calculating c , five separate samples were extracted for each random, aspect and $\cos i$ sampling method. The mean and standard deviation of the c -parameters resulting from the five trials were calculated for each sampling method and spectral band. If the standard deviation of c from the five samples using a single sampling method was large (i.e., the five c were widely different), that sampling method had a strong possibility of deriving inconsistent and potentially less reliable c .

3.3.2 Assessing accuracy of the c -parameter

There is no single standard method used to assess the success of a topographic normalization, and a combination of assessments is often recommended (Riaño et al. 2003), but the methods are not agreed upon (e.g., Bishop and Colby 2002). In this study we have used a combination of visual assessment, classification accuracy and two-sample independent t -tests to determine the effectiveness of reducing spectral variation due to topography. Visual assessment of the images before and after topographic normalization often gives an indication of the normalization's effect, but is subjective, and further quantitative methods should be used. One such method is to compare classification accuracies from pre- and post-topographically normalized images. In this study, the normalized images were classified with a maximum-likelihood algorithm using all wavelength bands, as removing any band gave lower classification accuracy. The independent evaluation dataset described in Section 2.4 was used to assess the classification accuracy. As an additional comparison, we assessed the effect on classification accuracy when applying a large range of c -parameters ($c = 0.1, 0.2, 0.3 \dots 2.0$) to each band. Another quantitative evaluation method is to observe the effect that normalization has on reducing the spectral variation over the full range of $\cos i$ (Civco 1989; Riaño et al. 2003). Using the evaluation dataset, two-sample independent t -tests were conducted to compare mean spectral values for low and high $\cos i$ groups (i.e., horizontal $\cos i \pm 0.15$ separated the two groups) for the individual alpine vegetation classes in Table 2. The t -tests were performed using an incremental range of c ($c = 0.1, 0.2, 0.3 \dots 2.0$) for each spectral band. The motivation for using the t -test was that it would be less sensitive to outliers in the sample than the more commonly used standard deviation (e.g., Riaño et al. 2003). Using a 95% confidence interval, the null hypothesis was that the normalized spectral mean values from the low and high $\cos i$ groups were equal to each other. The conditions indicating the least likelihood of rejecting the null

hypothesis that the groups were similar (i.e., when the t -value was large and p -value was low), indicated the c -parameter that best normalized the data for that vegetation class. Scatter plots for each alpine vegetation class were also useful for visualizing the effect of different c in the normalization.

4 Results

4.1 Precision of the c -parameter

The mean value of c and standard deviation resulting from the five sample trials are reported for each sampling method and image (Table 5). Of the different sampling methods, the *cosi* sample gave c -parameters with the highest precision (i.e., lowest standard deviation from the five trials), for all images and most bands. The random sample with $n = 16,500$ and the aspect sample also provided reasonably precise c -parameters, although the standard deviations were slightly higher (with a few exceptions) than those from the *cosi* sample. As was expected, the random samples with different n showed that as sample size decreased, the precision of c also decreased, with the lowest precision (highest standard deviation) from the $n = 1,600$ random sample. The *cosi* sample produced a larger value of c for the NIR band as compared to the other sampling methods, in particular for the SPOT 2008 and Landsat 2006 images (Table 5). For other wavelength bands, the *cosi* sample produced a slightly lower c for the visible and SWIR2 (TM band 7) bands, while for the SWIR1 band (TM band 5 and SPOT band 4) c was often similar between sampling methods. The mean coefficients of determination (R^2) from the linear regressions between *cosi* and the spectral data were much higher from the *cosi* sample (often at least double) than the other sampling methods, for all bands and all images (Table 6). Scatter plots of the distribution of spectral values across *cosi* from the different sampling methods are shown in Fig. 5, using the SPOT 2008 data as an example.

Table 5 The mean c -parameter and the standard deviation of c from five repeated trials for each sampling method. The lowest standard deviation (i.e., highest precision) is shown in **bold**. Random refers to the random sampling method, Aspect refers to the north-south aspect stratified sample, and *Cosi* refers to the cosine of i stratified sample with power allocation.

Image and sample	Mean c -parameter						Standard deviation					
	Blue	Green	Red	NIR	SWIR 1	SWIR 2	Blue	Green	Red	NIR	SWIR 1	SWIR 2
TM 2006												
Random $n=16,500$	0.882	0.880	0.452	0.517	0.240	0.138	0.019	0.018	0.015	0.015	0.006	0.006
Random $n = 5,000$	0.900	0.881	0.461	0.515	0.241	0.147	0.061	0.036	0.027	0.029	0.015	0.026
Random $n = 1,600$	0.787	0.817	0.399	0.516	0.215	0.111	0.086	0.059	0.033	0.042	0.023	0.024
Aspect $n = 5,000$	0.770	0.801	0.395	0.502	0.236	0.126	0.024	0.033	0.008	0.012	0.008	0.008
<i>Cosi</i> $n = 5,000$	0.709	0.773	0.357	0.729	0.240	0.079	0.009	0.010	0.014	0.010	0.003	0.005
TM 2005												
Random $n=16,500$	1.594	0.988	0.705	0.431	0.470	0.510	0.116	0.010	0.029	0.022	0.002	0.028
Random $n = 5,000$	1.710	1.020	0.741	0.421	0.491	0.572	0.280	0.064	0.058	0.015	0.015	0.030
Random $n = 1,600$	1.908	1.093	0.827	0.443	0.505	0.604	0.564	0.112	0.182	0.104	0.046	0.178
Aspect $n = 5,000$	1.347	0.889	0.615	0.456	0.438	0.440	0.145	0.055	0.050	0.008	0.028	0.045
<i>Cosi</i> $n = 5,000$	1.377	0.832	0.493	0.468	0.370	0.314	0.035	0.016	0.006	0.007	0.009	0.011
SPOT 2008												
Random $n=16,500$	--	0.747	0.361	0.252	0.299	--	--	0.028	0.020	0.007	0.011	--
Random $n = 5,000$	--	0.726	0.334	0.267	0.289	--	--	0.024	0.011	0.035	0.010	--
Random $n = 1,600$	--	0.746	0.351	0.270	0.306	--	--	0.055	0.024	0.007	0.014	--
Aspect $n = 5,000$	--	0.693	0.340	0.254	0.290	--	--	0.015	0.004	0.013	0.016	--
<i>Cosi</i> $n = 5,000$	--	0.687	0.319	0.421	0.314	--	--	0.013	0.005	0.003	0.003	--
SPOT 2004												
Random $n=16,500$	--	1.633	0.936	0.629	0.522	--	--	0.039	0.029	0.033	0.015	--
Random $n = 5,000$	--	1.725	0.999	0.618	0.537	--	--	0.119	0.093	0.049	0.029	--
Random $n = 1,600$	--	1.694	0.992	0.646	0.545	--	--	0.305	0.276	0.053	0.101	--
Aspect $n = 5,000$	--	2.285	1.925	0.531	0.788	--	--	0.170	0.226	0.038	0.029	--
<i>Cosi</i> $n = 5,000$	--	1.414	0.798	0.644	0.437	--	--	0.011	0.023	0.020	0.012	--

Table 6

The mean coefficient of determination (R^2) from the linear regression between $\text{cos } i$ and each spectral band, from five repeated trials for each sampling method and image. The slope and gradient of the linear regression are used to calculate c (Table 1). The highest R^2 are shown in **bold**. Random refers to the random sampling method, Aspect refers to the north-south aspect stratified sample, and $\text{Cos } i$ refers to the cosine of i stratified sample with power allocation.

Image and Sample TM 2006	R^2 of linear regression between $\text{cos } i$ and					
	Blue	Green	Red	NIR	Swir 1	Swir 2
Random $n=16,500$	0.116	0.289	0.201	0.314	0.445	0.218
Random $n = 5,000$	0.116	0.293	0.204	0.315	0.452	0.220
Random $n = 1,600$	0.133	0.313	0.222	0.322	0.470	0.240
Aspect $n = 5,000$	0.137	0.318	0.231	0.335	0.499	0.249
$\text{Cos } i$ $n = 5,000$	0.388	0.648	0.557	0.498	0.797	0.587
TM 2005						
Random $n=16,500$	0.027	0.145	0.083	0.344	0.338	0.123
Random $n = 5,000$	0.027	0.146	0.082	0.354	0.337	0.116
Random $n = 1,600$	0.024	0.145	0.075	0.348	0.324	0.109
Aspect $n = 5,000$	0.040	0.181	0.109	0.368	0.393	0.159
$\text{Cos } i$ $n = 5,000$	0.122	0.481	0.387	0.561	0.754	0.464
SPOT 2008						
Random $n=16,500$	--	0.323	0.281	0.401	0.512	--
Random $n = 5,000$	--	0.333	0.293	0.396	0.513	--
Random $n = 1,600$	--	0.336	0.290	0.399	0.517	--
Aspect $n = 5,000$	--	0.364	0.317	0.400	0.561	--
$\text{Cos } i$ $n = 5,000$	--	0.634	0.604	0.506	0.781	--
SPOT 2004						
Random $n=16,500$	--	0.114	0.089	0.241	0.262	--
Random $n = 5,000$	--	0.107	0.082	0.242	0.258	--
Random $n = 1,600$	--	0.116	0.091	0.253	0.264	--
Aspect $n = 5,000$	--	0.084	0.040	0.305	0.201	--
$\text{Cos } i$ $n = 5,000$	--	0.419	0.378	0.515	0.708	--

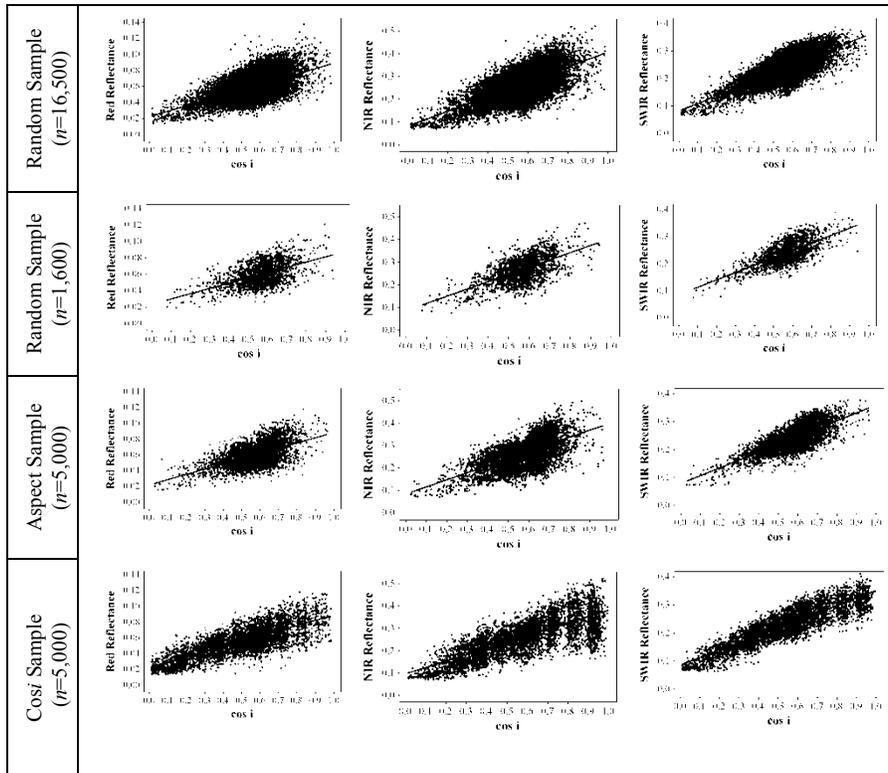


Fig. 5. Scatterplots of spectral reflectance (y-axis) versus cosine of i (x-axis) for four samples taken from the SPOT 2008 image. Columns are the red, near-infrared, and short-wave infrared bands. Different samples are given in the rows, with random sample ($n=16,500$) in first row, random sample ($n=1,600$) in second row, aspect sample ($n=5,000$) in third row, and $\text{cos}i$ sample ($n=5,000$) in fourth row. The line represents the regression line from which c is calculated (Eq. 3).

4.2 Accuracy of the c -parameter

The effect of topographic normalization using the c -parameters as determined by each sampling method was assessed using visual comparison, classification accuracy and t -tests. The results presented here are from the random sample ($n = 16,500$) and the *cosi* sample only, as the $n = 16,500$ random sample had c -parameters similar to the aspect sample, and produced c -parameters second best in precision to the *cosi* sample. Visually comparing the images normalized using c -parameters from the *cosi* sample and the random sample ($n = 16,500$) did not reveal large differences (Fig. 6), despite the difference in c -parameters. Over-correction of extreme slopes was not seen, and both normalized images appeared reasonable. Using a more quantitative evaluation method, classification accuracy for pre- and post-topographic normalization was assessed. For all four satellite images, slightly higher classification accuracies were obtained when the c -parameters derived from the *cosi* sample were used (Table 7). The resulting c -parameters differed most for the NIR band, and we therefore present results regarding the normalizing effect and accuracy of the c -parameters determined for the NIR band. Figure 7 shows the effect on classification accuracy when an incremental range of c was used to correct the NIR band, while the other bands' c -parameters were those derived by each respective sampling method and remained unchanged. The larger NIR c -parameter determined from the *cosi* sample resulted in higher classification accuracy. Figure 7 also shows that very low NIR c -parameters resulted in a classification accuracy result that was lower than if no topographic correction were performed at all; the very low c tended to overcorrect the spectral data.

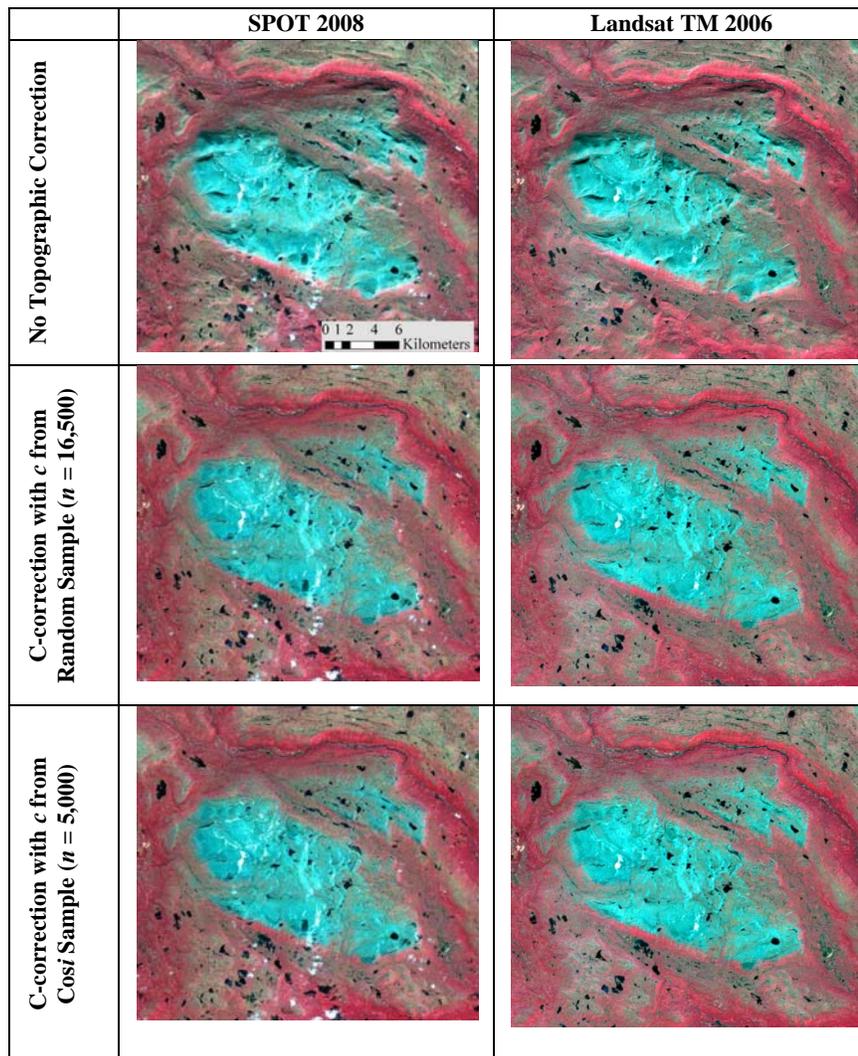


Fig. 6. Image subsets from SPOT 2008 (left column) and Landsat TM 2006 (right column). The first row is without topographic correction, the second row is C-correction using c -parameters from random sample ($n = 16,500$), and the third row is C-correction using c -parameters from *cosi* sample. Displayed bands: NIR, Red, Green (RGB).

Table 7

Classification accuracy for alpine vegetation types (listed in Table 2) with no topographic correction, using C-correction with c derived from the random sample ($n = 16,500$), and from the *cosi* sample ($n = 5,000$). Accuracy was assessed using the evaluation dataset ($n = 693$). All image bands were used in the classification. (Note: While the accuracies presented are low, classification methods for such detailed alpine vegetation classes often involve refinement of training data and use of ancillary data from a DEM to increase final map accuracy, which was not done for results presented here.)

	No topographic correction	C-correction random sample ($n = 16,500$)	C-correction <i>cosi</i> sample ($n = 5,000$)
Landsat 2006	47.4%	49.8%	50.5%
Landsat 2005	46.9%	49.0%	49.2%
SPOT 2008	51.4%	52.2%	53.3%
SPOT 2004	48.3%	49.6%	50.4%

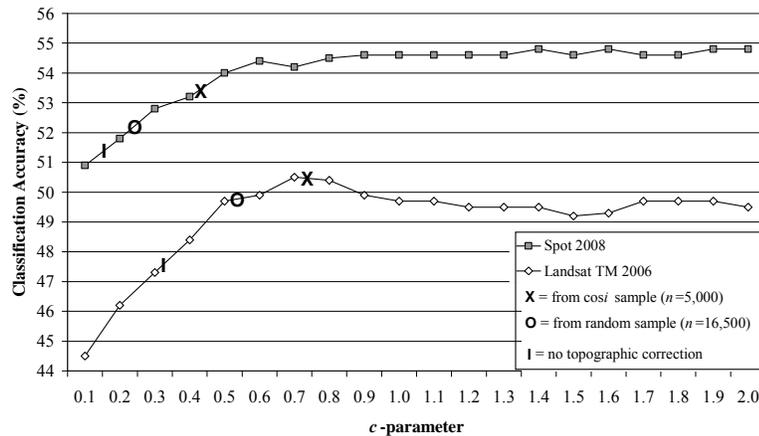


Fig. 7. The effect of increasing the NIR band's c -parameter on classification accuracy (all bands used), as assessed using the evaluation dataset ($n = 693$). The dark squares are from the SPOT 2008 image, the hollow diamonds are from the Landsat TM 2006 image. The classification accuracy and c -parameter as determined by the *cosi* sample is indicated by **X**, whereas **O** indicates the $n = 16,500$ random sample, and **I** indicates the classification accuracy if no topographic correction were to be performed. Landsat 2005 and SPOT 2004 images closely followed the curve for Landsat 2006, obscuring the results, and therefore are not presented here.

The two-sample independent t -tests (Table 8) indicated that optimal c -parameters differed among the individual alpine vegetation classes, with some general patterns seen among the four images (with exceptions). For the NIR band, dry alpine heath and alpine grass heath tended to have relatively higher c -parameters, mesic alpine heath was slightly lower, and alpine willow had relatively lower c -parameters, as did the alpine meadow classes. For these individual vegetation classes, the effect of normalizing the NIR band using c derived by the $n = 16,500$ random sample and the cosi sample is shown in scatter plots (Fig. 8).

Table 8 The optimal c -parameter for the NIR band as determined by the two-sample independent t -test for individual alpine vegetation classes are shown. Using the evaluation dataset (Table 2), the t -test compared the normalized mean spectral values between groups of low and high cosine of i (groups formed by horizontal $\text{cosi} \pm 0.15$), for each individual vegetation class. Tall alpine meadow and extremely dry alpine heath are not presented, because sufficient low cosi observations were lacking. The overall NIR c -parameter (based on all alpine vegetation types) from the random sample ($n = 16,500$) and the cosi sample for each image are also shown.

	Alpine dry heath NIR c	Alpine grass heath NIR c	Alpine mesic heath NIR c	Alpine short meadow NIR c	Alpine willow NIR c	Random sample $n=16,500$ NIR c	Cos <i>i</i> sample $n=5,000$ NIR c
Landsat 2006	0.7	0.7	0.7	0.3	0.4	0.517	0.729
Landsat 2005	1.2	0.8	0.7	0.2	0.3	0.431	0.468
SPOT 2008	0.6	0.7	0.5	0.4	0.4	0.252	0.421
SPOT 2004	1.0	1.3	0.8	0.8	0.5	0.629	0.644

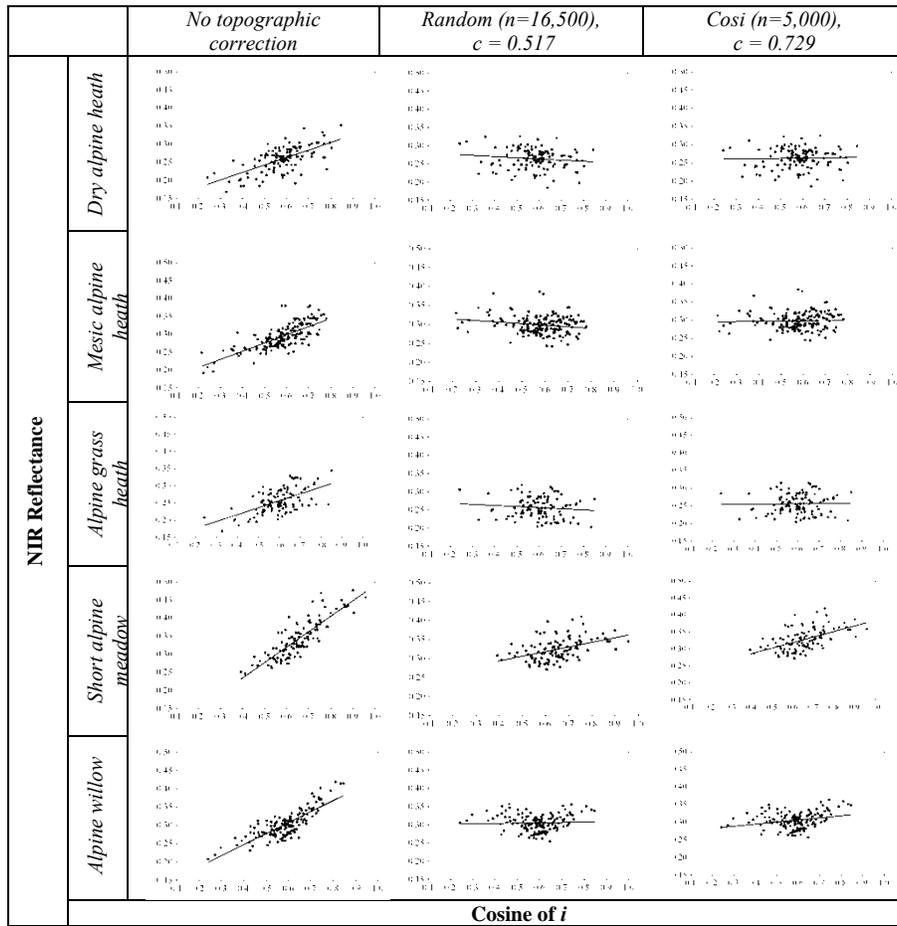


Fig 8. Scatterplots using the evaluation dataset of NIR reflectance (y-axis) versus cosine of i (x-axis) for individual alpine vegetation classes (in rows) from the Landsat TM 2006 image. Data without topographic correction (left column), C-correction of the NIR band using c as calculated by the $n = 16,500$ random sample (middle column), and C-correction of the NIR band using c as calculated by the $cosi$ sample (right column), are shown. When the line is horizontal, this shows the spectral data have been normalized over the range of cosine of i .

Image statistics (Table 4) as well as the scatter plots from the *cosi* sample (Fig. 5) indicated that the NIR reflectance had a larger standard deviation in the higher *cosi* strata than in the lower *cosi* strata. Unequal variance in the sample can complicate calculation of the empirical parameter. To investigate whether the sample was affected by the vegetation classes or whether the unequal variance was more an effect of the topographic and illumination angles, the NIR reflectance from the alpine vegetation subset was compared to another vegetation type, the sub-alpine birch forest subset (Fig 9). The NIR reflectance within the sub-alpine birch forest subset showed less variation across *cosi* strata than did the alpine vegetation subset. To gain more insight into the role that different alpine vegetation classes might play within the *cosi* sample, more detailed vegetation class labels were extracted from the GSD-Land Cover and Vegetation map (Fig. 10) for each sample. From this, it could be seen that grass heath and sparse vegetation had a different linear relationship within the sample than did dry/mesic heath, willow and alpine meadow.

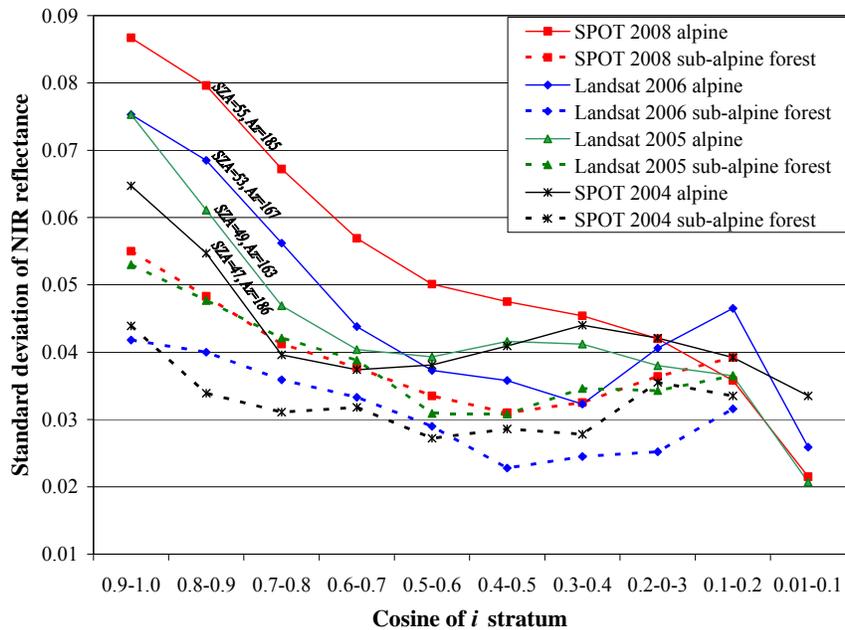


Fig. 9. Standard deviation of NIR reflectance (y-axis) for all pixels within each cosine of *i* stratum (x-axis) from both the alpine vegetation subset (solid line), and the sub-alpine forest subset (dashed line), for all four images. The solar zenith angle (SZA) and solar azimuth (Az) for each image is shown just above their respective data line.

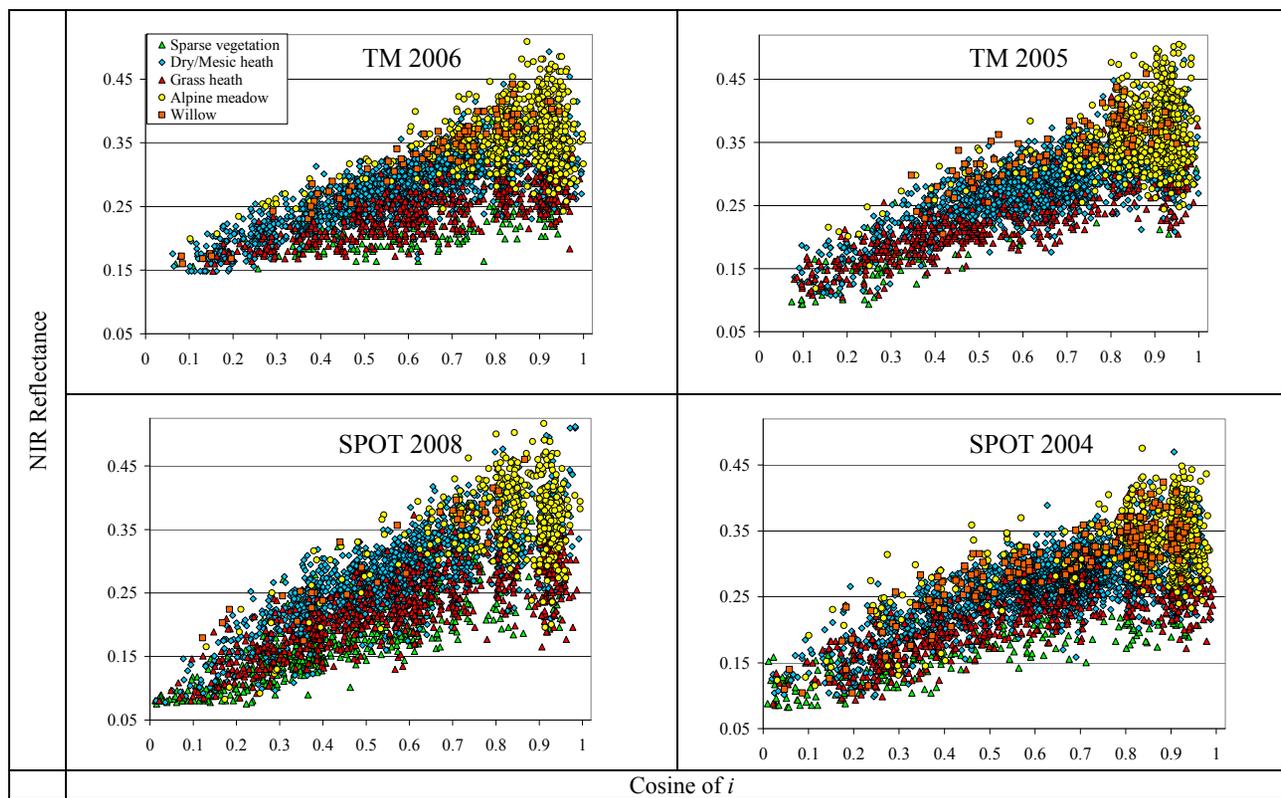


Fig. 10. NIR reflectance (y-axis) versus cosine of i (x-axis) from the $\text{cos}i$ sample ($n = 5,000$) drawn from each respective satellite image. Observations are labeled with the corresponding vegetation class taken from the GSD-Land and Vegetation Cover data.

5 Discussion

5.1 Precision of c

Of the five sampling methods compared in this study, the *cosi* sample most often produced the most precise c over repeated trials for all images (Table 5), while the random samples with $n = 5,000$ and $n = 1,600$ produced the least precise c -parameters. Since c is calculated by dividing the linear regression's intercept by its gradient (Eq. 3), the accurate determination of both intercept and gradient are important. In the study area, *cosi* (the independent variable) followed a normal distribution within the scene as a whole, and also within the alpine vegetation subset. A random sample of observations, although narrowed down to represent the alpine vegetation land cover only, also followed the normal distribution of *cosi*. The majority of observations were therefore clustered around the *cosi* mean, while observations for extreme values of *cosi* were sparse and influential. This rendered precise determination of the regression's intercept and gradient difficult and sometimes instable. Such relationships may lead to calculation of unsuitably large or small c -parameters, especially when the sample size is too small. This could be seen particularly in the $n = 1,600$ random sample (Fig. 5). While the $n = 16,500$ random sample had better representation, making c more precise, the representation for the high *cosi* values may still not have been sufficient to produce a correct c when compared to the spectral variation actually present in the data. In this study, more observations were needed from the extreme *cosi* values, giving stratification of the sample by *cosi* its advantage over random sampling. Although the aspect sample was also a stratified sample, in this study area it did not guarantee acquiring sufficient observations from low and high *cosi*, resulting in less reliable and less precise c -parameter calculation as compared to the *cosi* sample.

The advantage of stratifying the sample by *cosi* was also seen in the consistently stronger linear relationships (Table 6). A higher R^2 did not indicate

that the resulting c -parameter was correct, but that it was based on a dataset which better described the relationship between the variables and that would produce more precise estimates of c over repeated trials. The random samples had the weakest linear relationships, regardless of the sample size, due to the observations following the normal distribution of $\cos i$ in the study area. For any sampling method, the relatively low R^2 of the blue band was likely due to atmospheric scattering, while the consistently higher R^2 of the SWIR1 band may have been related to the reflectance characteristics of the alpine vegetation classes in the study area.

5.2 The advantage of power allocation in the cosine of i stratified sample

While stratification of the sample by $\cos i$ was advantageous, the particular allocation of the observations to each $\cos i$ stratum was also important. Fig. 5 shows that the $\cos i$ sample with $n = 5,000$ captured more spectral variability than even the largest random sample ($n = 16,500$), in particular for the higher $\cos i$ strata. When determining observation allocation in the stratified sample, consideration needs to be taken not only to the independent variable's ($\cos i$) distribution, but also to the dependent variable's (spectral data) distribution. The power allocation fulfills this purpose by incorporating statistics regarding the dependent variable. In addition, the power allocation is not dependent on stratum size, and can increase allocation within small but important strata. Since the power allocation is empirically-based and flexible, it should also be applicable for conditions and image data different from those presented in this study.

5.3 Accuracy of the c -parameter

The c -parameters resulting from the three sampling methods differed, with the largest difference being for the NIR band, where c was consistently larger from the $\cos i$ sample than the other sampling methods. The classification accuracy using c -parameters from the $\cos i$ sample was slightly higher than that from the $n = 16,500$ random sample, for all images. The reasons for the difference in classification accuracy could be partially explained by the evidence that optimal c -parameters differed depending on individual vegetation classes, as shown in the t -tests, as well as by considering the characteristics of the evaluation dataset used for accuracy assessment. Taking the Landsat 2006 image as an example, for three vegetation classes a higher NIR c ($c > 0.5$) was optimal (representing 390 of 693 evaluation plots), while for two classes a

lower NIR c ($c \leq 0.5$) was optimal (representing 261 of 693 evaluation plots). Therefore, a larger percentage of the evaluation plots were optimally corrected using the *cosi* sample's c -parameters, with the NIR $c > 0.5$. These results indicate that the *cosi* sample has produced appropriate c -parameters. However, the implication from the t -tests, that multiple vegetation types requiring different c -parameters were present in the sample, makes interpretation of the accuracy of c less straightforward. In actuality, if multiple vegetation types are present in a sample, this may lead to production of an “average c ” for those vegetation types, depending on their reflectance characteristics and quantities within the sample. Even with multiple vegetation classes present, however, calculating c from a sample stratified by *cosi* provides a better chance of calculating appropriate c -parameters due to the better data representation than from a random sample.

5.4 Importance of vegetation type specific normalization

Unequal variance in the data (in addition to non-linearity) is an issue occasionally reported in other topographic normalization studies, and in some cases has been addressed by linearizing the spectral data (Vincini and Frazzi 2003). The cause of unequal variance and non-linearity should be identified, however. If the root of these confounding factors is that the sample contains vegetation types with different reflectance behavior, then ideally, they should be separated into different samples. Due to the complex nature of anisotropic reflectance, it is difficult to pinpoint any single explanation regarding the NIR reflectance behavior in the high *cosi* strata seen in this study. It may be a result of multiple, spectrally different vegetation classes (some of whose ecological niches may be highly correlated with topographic characteristics), as well as the anisotropic behavior of NIR reflectance as affected by illumination angles, effective view angles, and the vegetation canopy (Goodin et al. 2004; Huete et al. 1992; Pinter et al. 1990; Qi et al. 1995; Vierling et al. 1997). Addressing the circular nature of vegetation class specific topographic normalization was outside the scope of this paper. However, the results indicated that calculation of empirical parameters for specific alpine vegetation types may be a subject worthy of future study. In particular, the classes with shorter and sparser vegetation (e.g., dry alpine heath, alpine grass heath) may require separation from taller and denser vegetation classes (e.g., mesic alpine heath, alpine willow), as well as the alpine meadow. Background soil reflectance may be an important contributing factor to the observed radiance for some of the alpine vegetation types with sparse canopy cover. Pinter et al. (1990) and Huete et al. (1992) determined that the amount of soil viewed through a vegetation canopy

contributes to the surface anisotropy. In this case, deriving separate normalizations for image subsets based on fraction of soil and green vegetation (Montandon and Small 2008) may be of interest. Vegetation indices such as NDVI [e.g., Bishop, 2002 #62] have been used for separating vegetation types, although other indices could be tested (e.g., Soil Adjusted Vegetation Index, SAVI), with the caution that ratio-based indices still contain topographic and illumination effects (Goodin et al. 2004; Huete et al. 1992; Veraverbeke et al. 2010; Verrelst et al. 2008).

5.5 Results for Landsat and SPOT

Four satellite images were used in this study, which were too few to draw conclusions regarding the effects of solar illumination angles upon the observed results. Issues related to SPOT's off-nadir view angle were also unaddressed in this study. It is, however, interesting that the SPOT 2008 image with the largest solar zenith angle (SZA=55°) had the highest NIR standard deviation within high *cosi* strata, while the SPOT 2004 image with the lowest solar zenith angle (SZA = 47°) had the lowest NIR standard deviation (Fig. 9). Increased NIR reflectance with increasing solar zenith angles and view angles has been reported in previous studies for grasslands in flat terrain (Huete et al. 1992; Pinter et al. 1990) and sloping terrain (Goodin et al. 2004). These same studies also showed that red reflectance increased with decreasing solar zenith angles, as the illumination of soil relative to grass increased.

The Landsat TM 2006 image (SZA = 53°) was most affected by completely shadowed pixels, which may have been due to the large solar zenith angle, while the SPOT 2008 image (SZA = 55°) was less affected by completely shadowed pixels, perhaps due to the sensor's eastward pointing view angle (+26.8°). In general, however, the *cosi* stratified sampling worked well for calculating *c* for all four images used here, despite their different illumination and view angles.

5.6 Summary of suggested method to determine the empirical parameter

From this study, we present suggestions regarding sampling for the calculation of appropriate *c*-parameters for C-correction. An image subset of the vegetation type to be normalized should be created, from which the empirical parameter will be estimated. We used, for example, an existing land cover map to identify a broad "alpine vegetation" subset. Any classes differing greatly in reflectance behavior should be excluded from the subset if possible, as done in

this study, where bare rock and snow were omitted using an NDVI threshold. To determine the minimum number of sample observations for calculation of the empirical parameter, sampling theory (e.g., Thigpen 1987) can help provide a guideline. The sample used to derive the empirical parameter should be stratified on $\cos i$, and in this study we used a 0.1 increment for the $\cos i$ strata. The allocation of observations to each $\cos i$ stratum can be determined using power allocation (Bankier 1988) which takes the spectral data properties into account and manipulates the influence of stratum size. We used $q = 0.3$ for all strata in the visible bands, while for the NIR and SWIR bands more weight was given to the higher $\cos i$ stratum ($q = 0.4$), and less to the lower $\cos i$ stratum ($q = 0.2$). The power allocation is flexible, however, and observation allocation can be determined according to the data characteristics of each particular study. A minimum threshold of the NIR reflectance based on negative $\cos i$ values can eliminate erroneously included observations of complete shadow from the sample. Using this sample, the c -parameters can be determined and the C-correction applied to the corresponding image subset. This strategy for determining the empirical c -parameter -- based on a sample stratified by cosine of i and with a power allocation of observations -- may also be useful in determination of other empirical parameters such as the Minnaert constant k (Smith et al. 1980), the b of b-correction (Vincini and Frazzi 2003), and C in SCS+C (Soenen et al. 2005).

6 Conclusions

By comparing the values of the C-correction's empirical parameter, c , as derived from three different sampling strategies (a random sample, a sample stratified on aspect, and a sample stratified on cosine of i), it could be seen that the calculation of c was influenced by the sample from which it was derived. The sample stratified by cosine of i (the "cosi sample") produced the most precise c -parameters (lowest standard deviation over five trials), and had the highest coefficients of determinations (R^2). A power allocation was used in the *cosi* sample, which provided better representation of the spectral variability within high cosine of i strata. This was particularly important for the NIR band. The increased representation of the spectral variability was the reason that the *cosi* sample produced different c -parameters as compared to the random and aspect samples. "Large" random samples ($n = 16,500$), taken for alpine vegetation only, did not represent the full spectral variability required to produce appropriate c -parameters for all alpine vegetation types. Evaluations (t -tests and classification accuracy) using photo-interpreted plots showed that the cosine of i sample's c -parameters corrected a larger portion of the evaluation data correctly as compared to the other samples.

Characteristics of the study area's dataset, such as the distribution of cosine of i , the spectral variability, and the vegetation types all exert influences that should be considered when taking the sample to derive c . Although the samples were narrowed down to alpine vegetation type only, the individual alpine vegetation classes within the subset may have benefited from differing c -parameters to achieve optimal topographic normalization. In particular, sparse vegetation types seemed to differ from denser growth types. The work in this study focused on non-forest alpine vegetation types using Landsat and SPOT images with relatively large solar zenith angles; the applicability of using a cosine of i stratified sample as described in this paper to calculate c for other vegetation types, or for the calculation of other empirical parameters, is a

subject for future study. As the sample's characteristics have been shown to influence the outcome of the empirical parameter, it's important that future topographic normalization studies provide information about the sample used to calculate empirical parameters.

Acknowledgments

The authors thank the Swedish National Space Board for funding of this project, as well as the anonymous reviewers who provided helpful comments on the manuscript.

References

- Bankier, M.D. (1988). Power allocations: Determining sample sizes for subnational areas. *The American Statistician*, 42, 174-177
- Bater, C.W., & Coops, N.C. (2009). Evaluating error associated with lidar-derived DEM interpolation. *Computers & Geosciences*, 35, 289-300
- Bishop, M.P., & Colby, J.D. (2002). Anisotropic reflectance correction of SPOT-3 HRV imagery. *International Journal of Remote Sensing*, 23, 2125-2131
- Blesius, L., & Weirich, F. (2005). The use of the Minnaert correction for land-cover classification in mountainous terrain. *International Journal of Remote Sensing*, 26, 3831-3851
- Carpenter, G.A., Gopal, S., Macomber, S., Martens, S., Woodcock, C.E., & Franklin, J. (1999). A neural network method for efficient vegetation mapping. *Remote Sensing of Environment*, 70, 326-338
- Civco, D.L. (1989). Topographic normalization of Landsat Thematic Mapper digital imagery. *Photogrammetric Engineering and Remote Sensing*, 55, 1303-1309
- Colby, J.D. (1991). Topographic normalization in rugged terrain. *Photogrammetric Engineering and Remote Sensing*, 57, 531-537
- Conese, C., Gilabert, M.A., Maselli, F., & Bottai, L. (1993). Topographic normalization of TM scenes through the use of an atmospheric correction method and digital terrain models. *Photogrammetric Engineering and Remote Sensing*, 59, 1745-1753
- Deng, Y.X., Chen, X.F., Chuvieco, E., Warner, T., & Wilson, J.P. (2007). Multi-scale linkages between topographic attributes and vegetation indices in a mountainous landscape. *Remote Sensing of Environment*, 111, 122-134
- Ekstrand, S. (1996). Landsat TM-based forest damage assessment: Correction for topographic effects. *Photogrammetric Engineering and Remote Sensing*, 62, 151-161
- Giardino, C., & Brivio, P.A. (2003). The application of a dedicated device to acquire bidirectional reflectance factors over natural surfaces. *International Journal of Remote Sensing*, 24, 2989-2995
- Giles, P.T. (2001). Remote sensing and cast shadows in mountainous terrain. *Photogrammetric Engineering and Remote Sensing*, 67, 833-839
- Goodin, D.G., Gao, J.C., & Henebry, G.M. (2004). The effect of solar illumination angle and sensor view angle on observed patterns of spatial structure in tallgrass prairie. *Ieee Transactions on Geoscience and Remote Sensing*, 42, 154-165

- Gu, D., & Gillespie, A. (1998). Topographic normalization of Landsat TM images of forest based on subpixel Sun-canopy-sensor geometry. *Remote Sensing of Environment*, *64*, 166-175
- Holben, B.N., & Justice, C.O. (1980). The topographic effect on spectral response from nadir-pointing sensors *Photogrammetric Engineering and Remote Sensing*, *46*, 1191-1200
- Huete, A.R., Hua, G., Qi, J., Chehbouni, A., & Vanleeuwen, W.J.D. (1992). Normalization of multidirectional Red and NIR reflectances with the SAVI *Remote Sensing of Environment*, *41*, 143-154
- Hugli, H., & Frei, W. (1983). Understanding anisotropic reflectance in mountainous terrain *Photogrammetric Engineering and Remote Sensing*, *49*, 671-683
- Kane, V.R., Gillespie, A.R., McGaughey, R., Lutz, J.A., Ceder, K., & Franklin, J.F. (2008). Interpretation and topographic compensation of conifer canopy self-shadowing. *Remote Sensing of Environment*, *112*, 3820-3832
- Kobayashi, S., & Sanga-Ngoie, K. (2009). A comparative study of radiometric correction methods for optical remote sensing imagery: the IRC vs. other image-based C-correction methods. *International Journal of Remote Sensing*, *30*, 285-314
- Lehtonen, R., & Pahkinen, E. (2004). *Practical Methods for Design and Analysis of Complex Surveys*. (2 ed.). Chichester: John Wiley & Sons
- Meyer, P., Itten, K.I., Kellenberger, T., Sandmeier, S., & Sandmeier, R. (1993). Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment *Isprs Journal of Photogrammetry and Remote Sensing*, *48*, 17-28
- Montandon, L.M., & Small, E.E. (2008). The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. *Remote Sensing of Environment*, *112*, 1835-1845
- Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, *97*, 558-625
- Nichol, J., Hang, L.K., & Sing, W.M. (2006). Empirical correction of low Sun angle images in steeply sloping terrain: a slope-matching technique. *International Journal of Remote Sensing*, *27*, 629-635
- Olthof, I., Butson, C., & Fraser, R. (2005). Signature extension through space for northern landcover classification: A comparison of radiometric correction methods. *Remote Sensing of Environment*, *95*, 290-302
- Pinter, P.J., Jackson, R.D., & Moran, M.S. (1990). Bidirectional reflectance factors of agricultural targets - A comparison of ground-based, aircraft-based, and satellite-based observations *Remote Sensing of Environment*, *32*, 215-228
- Proy, C., Tanre, D., & Deschamps, P.Y. (1989). Evaluation of topographic effects in remotely sensed data *Remote Sensing of Environment*, *30*, 21-32
- Qi, J., Moran, M.S., Cabot, F., & Dedieu, G. (1995). Normalization of sun/view angle effects using spectral albedo-based vegetation indexes *Remote Sensing of Environment*, *52*, 207-217
- Rafstedt, T. (1983). *Fjällensvegetation: Norrbottens län*. Stockholm, Sweden: Naturgeografiska Intitutet, Stockholm University. Liber distribution.
- Reese, H., Nilsson, M., & Olsson, H. (2009). Comparison of Resourcesat-1 AWiFS and SPOT-5 data over managed boreal forest stands. *International Journal of Remote Sensing*, *30*, 4957-4978

- Riaño, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types (2003). *Ieee Transactions on Geoscience and Remote Sensing*, *41*, 1056-1061
- Richter, R., Kellenberger, T., & Kaufmann, H. (2009). Comparison of topographic correction methods. *Remote Sensing*, *1*, 184-196
- Riedel, S.M., Epstein, H.E., & Walker, D.A. (2005). Biotic controls over spectral reflectance of arctic tundra vegetation. *International Journal of Remote Sensing*, *26*, 2391-2405
- Roberts, D.A., Smith, M.O., & Adams, J.B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data *Remote Sensing of Environment*, *44*, 255-269
- Roy, D.P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, *112*, 3112-3130
- Shepherd, J.D., & Dymond, J.R. (2003). Correcting satellite imagery for the variance of reflectance and illumination with topography. *International Journal of Remote Sensing*, *24*, 3503-3514
- Sibson, R. (1981). A brief description of natural neighbor interpolation. In V. Barnett (Ed.), *Interpolating Multivariate Data* (pp. 21-36). New York: John Wiley & Sons
- Smith, J.A., Lin, T.L., & Ranson, K.J. (1980). The Lambertian assumption and Landsat data *Photogrammetric Engineering and Remote Sensing*, *46*, 1183-1189
- Soenen, S.A., Peddle, D.R., & Coburn, C.A. (2005). SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *Ieee Transactions on Geoscience and Remote Sensing*, *43*, 2148-2159
- Soenen, S.A., Peddle, D.R., Coburn, C.A., Hall, R.J., & Hall, F.G. (2008). Improved topographic correction of forest image data using a 3-D canopy reflectance model in multiple forward mode. *International Journal of Remote Sensing*, *29*, 1007-1027
- Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Gallegos Torell, Å., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B., & Inghe, O. (2010). National Inventory of Landscapes in Sweden (NILS) - scope, design, and experiences from establishing a multiscale biodiversity monitoring system. *Environmental Monitoring and Assessment [Published online 17 March 2010]*
- Stow, D.A., Hope, A.S., & George, T.H. (1993). Reflectance characteristics of Arctic tundra vegetation from airborne radiometry. *International Journal of Remote Sensing*, *14*, 1239-1244
- Teillet, P.M. (1986). Image correction for radiometric effects in remote sensing. *International Journal of Remote Sensing*, *7*, 1637-1651
- Teillet, P.M., Guindon, B., & Goodenough, D.G. (1982). On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing*, *8*, 84-106
- Thigpen, C.C. (1987). A sample-size problem in simple linear regression. *The American Statistician*, *41*, 214-215
- Veraverbeke, S., Verstraeten, W.W., Lhermitte, S., & Goossens, R. (2010). Illumination effects on the differenced Normalized Burn Ratio's optimality for assessing fire severity. *International Journal of Applied Earth Observation and Geoinformation*, *12*, 60-70

- Verrelst, J., Schaepman, M.E., Koetz, B., & Kneubuhler, M. (2008). Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sensing of Environment*, 112, 2341-2353
- Vicente-Serrano, S.M., Perez-Cabello, F., & Lasanta, T. (2008). Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. *Remote Sensing of Environment*, 112, 3916-3934
- Vierling, L.A., Deering, D.W., & Eck, T.F. (1997). Differences in arctic tundra vegetation type and phenology as seen using bidirectional radiometry in the early growing season. *Remote Sensing of Environment*, 60, 71-82
- Vincini, M., & Frazzi, E. (2003). Multitemporal evaluation of topographic normalization methods on deciduous forest TM data. *Ieee Transactions on Geoscience and Remote Sensing*, 41, 2586-2590