

THE EFFECT OF STRATIFIED TOPOGRAPHIC CORRECTION ON LAND COVER CLASSIFICATION IN TROPICAL MOUNTAINOUS REGIONS

A. Twele*, M. Kappas, J. Lauer, S. Erasmí

University of Göttingen, Institute of Geography, Dept. of Cartography, GIS & Remote Sensing,
Goldschmidtstr. 5, 37077 Göttingen, Germany; atwele@uni-goettingen.de

KEY WORDS: Forestry, Land Cover, Classification, Correction, Algorithms, DEM/DTM

ABSTRACT:

Topographic correction is an important data preprocessing step when land cover classification and quantitative analysis of multispectral data are carried out in mountainous regions. Solar illumination effects may cause variations in reflectance of similar ground features, leading to a possible misclassification due to different topographic positions. Since nearly all surfaces exhibit a varying degree of anisotropy in bidirectional reflectance, the terrain normalization methodology should ideally be tailored to each vegetation type and account for the variability of land cover types within the research area. This study evaluated different topographic correction algorithms using statistical analysis and quantified their impact on classification accuracy. A pre-stratification of the dataset using continuous parameters was evaluated in order to model the degree of non-lambertian reflectance behaviour separately for the most prevalent vegetation types of the study area. The results suggest that the method used to extract semi-empirical constants has a strong influence on correction results and classification accuracy. The derivation of pixel-specific Minnaert constants based on a stratification according to terrain slope and vegetation index thresholds improved classification accuracies while not requiring *a priori* knowledge.

1. INTRODUCTION

Besides sensor variation and atmospheric effects, landscape topography can seriously affect the radiometric quality of remote sensing data. Topographically induced illumination effects may cause reflectance variations of similar ground features, leading to a possible misclassification into different land cover classes due to different topographic positions. The impact of these effects is mainly influenced by the local solar incidence angle (i.e. the angle between the incoming direct sun rays and the normal to the ground), the wavelength of incoming radiation and the specific reflectance characteristics of the land cover type. Various studies have shown that a correction of topographic effects that is carried out before multispectral and multitemporal pixel classification can substantially improve classification accuracies, particularly if the study site is located in rugged terrain (Meyer et al., 1993; Colby et al., 1998; Riaño et al., 2003).

Despite its benefits, the applicability of topographic normalization is often hampered through the unavailability of accurate digital elevation data, a problem that is especially prevalent in remote areas. This obstacle has been partially overcome through the completion of the Shuttle Radar Topography Mission (SRTM), providing nearly global digital terrain data. However, depending on the spatial resolution of optical remote sensing data, the spatial accuracy is still rather low and in need of improvement.

This study considered different topographic normalization algorithms (both lambertian and non-lambertian) and assessed their accuracy by statistical analysis and by comparing classification accuracies of different land cover types. The study site lies in the tropical rain-forest region of Central Sulawesi, Indonesia. All semi-empirical correction approaches should ideally account for the degree of lambertian to non-lambertian reflectance behaviour of each type of land cover

(Blesius and Weirich, 2005). Consequently, regarding the frequently used Minnaert correction, separate k-constants would ideally be calculated for all land cover types. However, this task is often problematic since the identification of land cover types is rather a processing step subsequent to topographic normalization. Thus, an appropriate method for pre-stratifying the area with respect to structural surface characteristics is required that ideally does not require *a priori* knowledge. Accordingly, a pre-stratification strategy based on continuous image-derived variables was evaluated in order to model the degree of non-lambertian reflectance behaviour of the dominant vegetation types of the study area.

2. STUDY AREA

The study area is located in Central Sulawesi, Indonesia, and encompasses the Lore Lindu National Park and five surrounding districts within the province of Central Sulawesi. The project area covers approximately 7.500 sq. kilometres (0°50'S – 2°04'S and 119°40'E – 120°30'E), whereof 2.200 sq. kilometres constitute the Lore Lindu National Park, a natural forest reserve that is facing remarkable forest conversion activities along the forest and park boundaries. Major parts of the park are characterized by steep slopes and a rugged relief with elevations ranging from sea-level up to 2600 meters. The mountain ranges are interspersed with narrow and outstretched valleys at different elevations and expositions. The central and southern parts of the research area are characterized by a humid tropical climate with mean annual temperatures between 25 and 26 °C at sea level, high humidity and a mean annual precipitation of more than 2500 mm that is subject to a high variability due to the diverse topography (Whitten et al., 2002). The annual precipitation distribution permits year-round agriculture. At the northern boarder of the Lore Lindu National Park, a sub-region of approximately 270 sq. kilometres was considered for land cover classification. The northern parts of

this test site comprise a lowland area where different types of anthropogenic land use (cocoa plantations, maize cultivation, paddy fields) are prevalent. Major parts of the mountain ranges which lie in the southern part of the sub-region are covered by relatively undisturbed tropical forest. Near the national park boarder however, many previously forested areas on sloping land have been transformed into areas of cocoa or maize cultivation. Some of these areas have been abandoned after short-term cultivation and reverted to bushland or secondary forest.

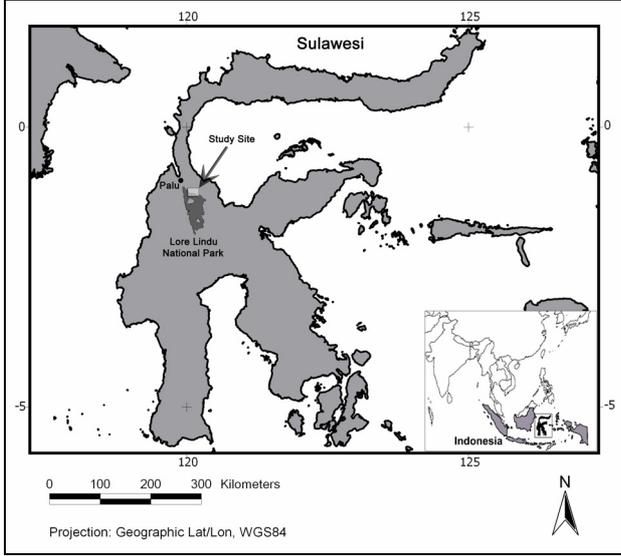


Figure 1. Location of the study site

3. TOPOGRAPHIC CORRECTION

The topographic correction algorithms which were evaluated in this study represent two basic categories, namely lambertian and non-lambertian correction approaches. The simplest model, i.e. the Cosine correction, assumes that incident radiation is reflected in all directions equally. This hypothesis of lambertian reflectance characteristics considers only direct irradiance although the proportion of diffuse skylight can be relatively high, particularly in weakly illuminated areas. Several non-lambertian methods have been developed to account for the shortcomings of the lambertian assumption. Although not truly physically based, these methods approximate bidirectional reflection characteristics by introducing empirically derived constants. These constants are used to model diffuse skylight and estimate the degree of non-lambertian behaviour of a given surface. The key factors influencing these constants are wavelength, phase angle and the type of surface cover (Meyer et al., 1993; Tokola et al., 2001).

The Cosine correction proposed by Teillet et al. (1982) only employs the solar zenith angle Θ_{sz} and the modelled terrain illumination $\cos(i)$ in the equation:

$$Ln = L \times \frac{\cos(\Theta_{sz})}{\cos(i)} \quad (1)$$

Although several authors have pinpointed the limitations of the lambertian approach (Meyer et al., 1993; Itten et al., 1992), the Cosine correction is probably still most widely used since it is incorporated as a standard procedure in most image processing

software. The Cosine correction tends to overcorrect topographic influences, which is most apparent when the algorithm is applied for the correction of steep and naturally vegetated surfaces. Despite its shortcomings, it was decided to include the Cosine correction in the evaluation to assess how the assumption of wavelength-independent lambertian reflectance characteristics would affect classification results.

Initially developed for photometric analysis of lunar surfaces (Minnaert, 1941), the Minnaert equation has found widespread application for the topographic normalization of optical satellite imagery (Teillet et al., 1982; Smith et al., 1980). The Minnaert constant k models the extent to which a surface has non-lambertian reflectance properties. Since the Minnaert constant is wavelength-dependent, separate constants need to be derived for each band. The term k is equal to the slope of the following regression equation:

$$\log(L \times \cos \Theta_n) = \log Ln + k \times \log(\cos(i) \times \cos(\Theta_n)) \quad (2)$$

The Minnaert constant k extends the Cosine correction as follows:

$$Ln = L \times \left(\frac{\cos(\Theta_{sz})}{\cos(i)} \right)^k \quad (3)$$

Colby (1991) included the cosine of the terrain slope as a separate term in the equation. For a nadir-viewing sensor such as Landsat, this angle is equivalent to the exitance angle.

$$Ln = L \times \cos(\Theta_n) \times \left(\frac{\cos(\Theta_{sz})}{\cos(i) \times \cos(\Theta_n)} \right)^k \quad (4)$$

The C-correction is based on a similar semi-empirical approach as the Minnaert correction. The parameter c is introduced as an additive term to counterbalance and prevent possible overcorrection which often occurs when using a solely lambertian approach such as the Cosine correction. The c -factor can be derived from the regression of $L = m \times \cos(i) + b$ as the quotient of the gradient and intercept of the linear regression equation, so that $c = \frac{b}{m}$. Terrain corrected reflectance Ln is then derived using the following equation, proposed by Teillet et al. (1982):

$$Ln = L \times \left(\frac{\cos(\Theta_{sz}) + c}{\cos(i) + c} \right) \quad (5)$$

4. DATA AND METHODS

4.1 Satellite Data

The base dataset consists of a Landsat ETM+ full scene (WRS-2, Path 114 / Row 61) of 28 September 2002. The scene was orthorectified using 65 well distributed ground control points (GCPs) that were derived from 1:50.000 scale maps and a local solar incidence angle map. The solar inclination map was advantageous for the detection of appropriate GCPs since an improved coregistration to the digital elevation model could be achieved which greatly minimized artefacts due to misregistration during topographic normalization.

In the humid tropics, atmospheric effects are particularly strong and can seriously affect the radiometric quality of optical satellite data. Since atmospheric impacts vary both spatially and temporally, there is a need to account for these effects especially within multi-temporal studies of land cover change. Although physically-based models which require atmospheric data coincident with remote sensing observations are considered to be most accurate (Lu et al., 2002), their overall applicability has been rather limited so far since these parameters are commonly not available for historical satellite data. Several image-based techniques have been developed in order to compensate for this shortcoming. We used the improved dark-object-subtraction method proposed by Chavez (1996) (also referred to as the COST-model) to account for the diametrical effects of both path radiance and atmospheric transmittance. The correction model was combined with sensor calibration, sun illumination and sun-earth distance normalization in a streamlined model to convert uncorrected digital numbers (DNs) to atmospherically corrected reflectance.

4.2 Digital Elevation Data

As for most remote areas, high-precision digital elevation data for Sulawesi is scarce and currently only available from an airborne SAR-interferometry mission. Due to high costs, this technique is only feasible for small parts of the project area. Although the generation of digital elevation models from optical stereo data or spaceborne SAR-interferometry would theoretically be less cost-constrained, these techniques have severe limitations over densely vegetated rain forest. Instead, the digital elevation data employed for the study was derived from digitized contour lines (equidistance of 25 metres) of topographic maps at a scale of 1:50,000. The digital elevation model (DEM) was resampled to 30 metre resolution to match the spatial grid cell size of the Landsat ETM+ sensor. In comparison to elevation data compiled from the SRTM (Shuttle Radar Topography Mission), the DEM showed increased topographical detail, particularly in steep mountainous areas. Several topographic variables were calculated from the digital elevation model and time- and site-specific solar information. In order to reconstruct illumination conditions from the time of imaging, the cosine of the angle between the surface normal and the direct solar beam was computed as follows:

$$\cos(i) = \cos \Theta_n \times \cos \Theta_{sz} + \sin \Theta_{sz} \times \sin \Theta_n \times \cos(\Phi_s - \Phi_n) \quad (6)$$

where

Θ_n	=	terrain slope angle
Θ_{sz}	=	solar zenith angle
Φ_s	=	solar azimuth angle
Φ_n	=	aspect angle

4.3 Methodology

Besides the correction algorithms mentioned in the previous chapter, the analysis included an uncorrected dataset as well as a dataset where sunlit and shadowed areas have been classified separately. The topographic correction algorithms were implemented in IDL and integrated into the ENVI GUI. Both non-lambertian algorithms automatically calculate Minnaert and C-constants, so no time-consuming manual extraction of pixel values and subsequent regression analysis is required. A disadvantage of this approach is that these constants are calculated globally from the entire scene, although previous work has shown that they can differ according to the type of

land cover (Bishop and Colby, 2002) and the density of vegetation cover (Twele and Erasmi, 2005). An additional problem is that the regression equation which is used to derive these empirical constants is negatively affected when major parts of the study region include areas of relatively flat terrain. This reduces the explanatory power of the equation and can lead to inappropriate constants that have a negative effect on correction accuracy. In order to circumvent these shortcomings, an additional stratified correction approach was implemented for the Minnaert-correction. The stratification is two-folded: First, only image pixels with a slope of more than 10 degrees are considered. Subsequently, the scene is subdivided into sub-regions that are determined by user-defined thresholds of continuous surface parameters. The thresholds were selected in such a way that each class contained the same amount of pixels. For our case study, the NDVI was used to stratify the uncorrected image into three classes for which separate Minnaert constants were calculated automatically.

The evaluation of correction algorithms was based on the computation of statistical parameters and by quantifying the classification accuracy of each method. Linear regression analysis between corrected data and the $\cos(i)$ was used to assess to what degree the dependency of reflectance on terrain illumination was reduced. Thus, the linear regression equation of an ideal correction algorithm would theoretically have a slope and R^2 of zero. The statistical analysis was based on a set of 5,000 randomly selected pixels. In this way, we were able to assess the degree of improvement gained by a stratified approach, where the surface characteristics and slope angle is accounted for in empirical sampling. Topographic correction should theoretically improve class homogeneity since illumination effects constitute a major source of class variability which increases misclassification. The classification of Landsat data was conducted using a standard maximum likelihood classifier. Class definitions were based on the LCCS (Land Cover Classification System, FAO; Di Gregorio & Jansen, 2000) with a total of eight land cover classes that were considered. Classification accuracy was assessed by calculating producer, user and overall accuracy as well as Kappa statistics. Individual class accuracies of critical land cover classes were checked in the cross tabulations statistics.

5. RESULTS

5.1 Empirical Constants

The method which was used for the automatic calculation of the empirical constants for both non-lambertian approaches proved to be very critical. An inclusion of all scene-pixels for the calculation of the constants resulted in substantially different constants than when a stratified approach was used for their calculation. The k-constants of the Minnaert correction were higher in all wavelengths when the pixels of the whole scene were considered. In two channels they approached, and in one they even surpassed values that are characteristic for lambertian reflectance ($k = 1$), which is unrealistic for densely vegetated surfaces in the optical domain. The three sets of surface-specific Minnaert constants that were extracted using the stratified approach were lower in all bands and resembled typical non-lambertian reflectance characteristics. Since the stratified sampling omitted areas with a slope of less than 10 degrees, only a subset of the total scene pixels was considered in the regression that is used to retrieve k-constants.

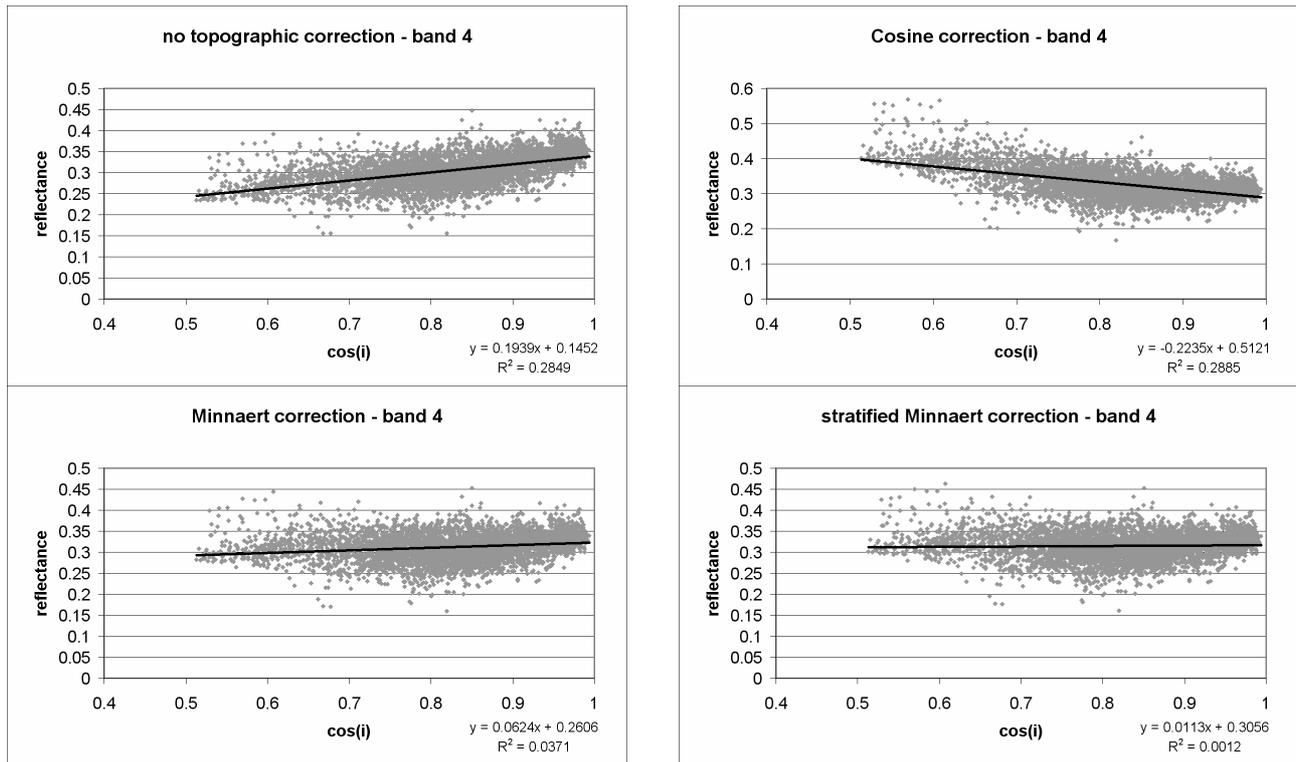


Figure 2: Statistical assessment of topographic correction accuracy

5.2 Statistical Analysis

The strong influence of terrain illumination on uncorrected reflectance is very evident in the study area, both visually and statistically. The topographic impact varies with wavelength and is most prevalent in the near-infrared spectral region (shown in Figure 2).

The Cosine correction showed the worst results of all corrections that have been evaluated. Especially slopes that face away from direct sunlight are obviously overcorrected which is very evident from the negative slope of the scatter plot (see Figure 2). The overcorrection is especially pronounced at lower ranges of $\cos(i)$ where the influence of diffuse skylight increases and direct solar irradiance is small. Instead of reducing the effect of $\cos(i)$ on surface reflectance, the Cosine correction actually increases it which renders the dataset even more distorted than before the correction.

In contrast, both non-lambertian methods (i.e. C-Correction and Minnaert correction) strongly reduced topographically induced illumination effects. Analogous to Meyer et al. (2003), we observed that only minor differences exist between the Minnaert and C-Correction. The substantial improvements achieved by using non-lambertian methods can primarily be explained by the reflectance characteristics of the predominant land cover of the test site (i.e. tropical rain forest), which strongly diverges from the lambertian assumption. This is underlined by Minnaert constants which lie far below 1 for the stratified approach in all bands, indicating rather non-lambertian reflectance properties. The stratification further lowered the slope (0.0113) and R^2 (0.0012) of the linear regression equation. This indicates an improvement in correction accuracy by a pixel-based instead of a global application of the Minnaert correction.

Areas where cast shadowing occurs were falsely corrected regardless of the correction methodology which was used. Although these areas do not receive any direct solar radiation due to shadowing from surrounding hills, the correction assumes the adverse since $\cos(i)$ is high. However, cast shadows only occur in areas where the terrain slope is extremely steep or sun elevation is very low.

5.3 Classification Accuracy

Concerning individual class accuracies, the largest difference of correction approaches occurred between the open and closed rain forest class. This is not surprising since open and closed rain forest dominates sloping land while being nearly completely replaced by different forms of anthropogenic land use in the flat areas. Consequently, these classes were most influenced by topographic correction while classification accuracies of land cover types that are typical of lowland areas (not listed in Table 1) did not differ significantly.

A separate classification of sunlit and shaded areas within each land cover type did not improve classification accuracy notably because the separability of critical classes (i.e. sunlit closed and shaded open forest) still remained problematic without normalizing for topographic influences before training sets were generated.

As a result of severe overcorrection, which was especially strong in the visual bands, the overall classification accuracy and Kappa statistic was lowest using the Cosine correction.

Both non-lambertian correction approaches (i.e. Minnaert- and C-Correction) significantly reduced mismatches which occurred in areas of open forest when no correction was applied. In general, a high level of agreement in classification results can be observed when comparing the Minnaert and C-Correction.

Table 1. Overview of classification accuracy

	Uncorrected	Traditional	Cosine	C-Correction	Minnaert	stratified Minnaert
Producer Accuracy (%)						
Closed Rain Forest	86.3	87.2	85.9	88.3	87.7	92.6
Open Rain Forest	69.0	70.3	49.5	74.3	74.8	79.2
Plantations (Cocoa/Coffee)	88.0	85.6	88.6	86.9	85.2	85.0
(...)	(...)	(...)	(...)	(...)	(...)	(...)
Overall accuracy (%)	78.9	79.2	74.2	80.0	79.5	83.2
Kappa statistic	0.726	0.729	0.659	0.737	0.730	0.780

However, the strategy that was used for the estimation of k-constants had a much higher impact on classification results. The stratified Minnaert correction showed the highest overall classification accuracy (83.2%) and best Kappa statistic (0.780). Thus, these results are congruent with the findings that were previously gained from a purely statistical evaluation of the methods. Individual class accuracy notably improved for both forest classes while the accuracy was slightly lower for plantations. Since the NDVI of plantations can reach levels that are equivalent to natural forests although vegetation structure is substantially different, the k-constant which were derived might not have been optimal for this land cover class.

6. CONCLUSION

The present study evaluated the impact of different topographic correction methods on image statistics and classification accuracies for a test site in a tropical mountainous region. A secondary target of this work was to assess the impact of different strategies for deriving k-constants that are employed by the Minnaert correction.

The simplistic assumption of lambertian surface properties, as presumed by the Cosine correction, is not appropriate for the terrain correction of tropical forest. We showed that forest canopies exhibit strong non-lambertian reflectance characteristics. Additional shortcomings of the Cosine correction are the negligence of diffuse solar flux and the missing consideration of the wavelength.

In contrast to the Cosine correction, all non-lambertian methods successfully reduced topographically induced illumination effects, which was confirmed visually, statistically and by an assessment of classification accuracy. Only minor differences exist between both non-lambertian methods which were considered in this study. Other factors such as the type of pre-stratification and the quality of empirical constants seem to have a more substantial impact on the accuracy of terrain normalization.

An analysis of the classification results showed that the method used for deriving and applying the k-factors is very critical since it strongly influences the accuracy of the correction. Accordingly, it is important to stress that all semi-empirical approaches need to account for the degree of lambertian to non-lambertian reflectance behaviour of each type of land cover. Thus, an appropriate method for pre-stratifying the area with respect to structural surface characteristics such as vegetation density is required. Stratification based on a minimum slope angle of 10 degrees and three NDVI classes proved to be suitable for this task and substantially improved classification

accuracy. However, further research and an application to other biomes are needed to verify the robustness of different image-derived parameters as a means to pre-stratify areas with respect to separate Minnaert constants. Unsupervised clustering techniques should be considered as an alternative approach to compute surface-specific constants without *a priori* knowledge. While this work only assessed the stratified correction approach for the Minnaert correction, it may also be valid for other non-lambertian methods such as the C-Correction.

REFERENCES

- Bishop, M. P., Colby, J. D., 2002. Anisotropic reflectance correction of SPOT-3 HRV imagery. *International Journal of Remote Sensing* 23(10), pp. 965-974.
- Blesius, L., Weirich, F., 2005. The use of the Minnaert correction for land-cover classification in mountainous terrain. *International Journal of Remote Sensing* 26(17), pp. 3831-3851.
- Chavez, P. S. Jr., 1996. Image-based atmospheric corrections - revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62, pp. 1025-1036.
- Colby, J. D., 1991. Topographic normalization in rugged terrain. *Photogrammetric Engineering and Remote Sensing* 57, pp. 531-537.
- Colby, J. D., Keating, P. L., 1998. Land cover classification using Landsat TM imagery in the tropical highlands: the influence of anisotropic reflectance. *International Journal of Remote Sensing* 19(8), pp. 1479-1500.
- Di Gregorio, A., Jansen, L. J. M., 2000. Land Cover Classification System (LCCS): Classification Concepts and User Manual. FAO, Rome.
- Itten, K. I., Meyer, P., Kellenberger, T., Leu, R., Sandmeier, S. T., Bitter, P., and Seidel, K., 1992. Correction of the impact of topography and atmosphere on Landsat TM forest mapping of Alpine regions. *Remote Sensing Series, Vol. 18*, University of Zurich.
- Lu, D., Mausel, P., Brondizio, E. and Moran, E., 2002. Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. *International Journal of Remote Sensing* 23(13), pp. 2651-2671.

Meyer, P., Itten, K. I., Kellenberger, J., Sandmeier, S. and Sammeier, R., 1993. Radiometric correction of topographically induced effects on Landsat TM data in an alpine environment. *ISPRS Journal of Photogrammetry and Remote Sensing* 48, pp. 17–28.

Minnaert, M., 1941. The reciprocity principle in lunar photometry. *Astrophysical Journal* 93, pp. 403–410.

Riaño, D., Chuvieco, E., Salas, F. J. and Aguado, I., 2003. Assessment of Different Topographic Corrections in Landsat-TM Data for Mapping Vegetation Types. *IEEE Transactions on Geoscience and Remote Sensing* 41, pp. 1056-1061.

Smith, J. A., Tzeu Lie Lin and Ranson, K. J., 1980. The lambertian assumption and Landsat data. *Photogrammetric Engineering and Remote Sensing* 46(9), pp. 1183-1189.

Teillet, P. M., Guindon, B., and Goodenough, D. G., 1982. On the Slope-Aspect Correction of Multispectral Scanner Data. *Canadian Journal of Remote Sensing* 8(2), pp. 84 - 106.

Tokola, T., Sarkeala, J., and van der Linden, M., 2001. Use of topographic correction in Landsat TM-based forest interpretation in Nepal. *International Journal of Remote Sensing* 22(4), pp. 551-563.

Twele, A., Erasmi, S., 2005. Optimierung der topographischen Normalisierung optischer Satellitendaten durch Einbeziehung von Kohärenzinformation. *Photogrammetrie Fernerkundung Geoinformation (PFG)* 3, pp. 227-234.

Whitten, A., Mustafa, M., and Henderson, G., 2002. The Ecology of Sulawesi. The Ecology of Indonesia Series, 4, Jakarta, Indonesia.

ACKNOWLEDGEMENTS

This work is funded by the German Research Foundation (DFG) as part of SFB-552 (“*STORMA* – Stability of Rainforest Margins in Indonesia”).