Can. J. Remote Sensing, Vol. 33, No. 4, pp. 325–340, 2007

A comparison of radiometric normalization methods when filling cloud gaps in Landsat imagery

E.H. Helmer and B. Ruefenacht

Abstract. Mapping persistently cloudy tropical landscapes with optical satellite imagery usually requires assembling the clear imagery from several dates. This study compares methods for normalizing image data when filling cloud gaps in Landsat imagery with imagery from other dates. Over a complex tropical island landscape, namely St. Kitts and Nevis and the island of St. Eustatius, all of the methods tested reduce interdate image differences for ETM+ bands 1–5 and 7, NDVI, and band 4:5 ratio. Regression tree normalization reduces the interdate differences more consistently than linear regression or histogram matching. Normalizing ETM+ images with regression trees can produce more seamless imagery than linear normalization, histogram matching, or image-based atmospheric correction via dark object subtraction. More seamless imagery enhances visual interpretation and helps reveal the distributions of forest formations in these landscapes. Decision tree classification of cloud-filled Landsat imagery can accurately map land cover and detailed forest formations. Decision tree classification accuracy is not highly dependent on the method used to make the cloud-filled imagery, however, at least as long as (*i*) classification model training data reflect class spectral variability, and (*ii*) ancillary spatial data that relate to the distributions of classes are used in the classification. Cloud-filled imagery is also known as cloud-cleared imagery.

Résumé. Les couvertures terrestres sont difficiles à cartographier à l'aide d'imagerie satellite à haute résolution lorsqu'il s'agit de paysages perpétuellement couverts de nuages. Cette étude compare les différentes méthodes utilisées pour normaliser les données d'images afin de remplir les trouées de nuages dans les images de Landsat avec des images d'autres dates. En balayant un paysage complexe d'une île tropicale, toutes les méthodes vérifiées réduisent les différences d'images entre les dates pour ce qui est des bandes 1 à 5, et 7 ETM+, de l'IVDN et du rapport de bandes 4:5. Pour ces bandes et indices, la normalisation de l'arbre de régression réduit ces différences plus que la régression linéaire, l'appariement d'histogrammes. Par conséquent, la normalisation d'images ETM+ à l'aide d'arbres de régression peut produire des images plus homogènes. Ce genre d'images facilite l'interprétation visuelle et aide à révéler la distribution des formations de forêts dans ces paysages. La classification basée sur un arbre de décision des images remplies de nuages de Landsat peut dresser avec précision la carte de la couverture terrestre et les formations de forêts en détail. La précision de la classification basée sur un arbre de décision dépend peu de la méthode utilisée pour obtenir les images remplies de nuages, cependant, du moins tant que (*i*) les données d'apprentissage du modèle de classification reflètent la variabilité spectrale de la classe et tant que (*ii*) les données spatiales auxiliaires liées aux distributions des classes sont utilisées dans la classification. On parle parfois de la notion d'enlèvement des nuages en imagerie de zones couvertes de nuages.

Introduction

Persistent cloud cover makes mapping some landscapes with optical satellite imagery seem daunting. Cloud-free imagery for one season can require assembling the cloud-free parts of several scene dates, and the seam lines from different dates are usually visible in the resulting image mosaics. This study addresses (*i*) potential methods for normalizing the Landsat images that fill clouds in a reference Landsat image, and (*ii*) classifying cloud-filled imagery over complex tropical landscapes with decision trees. Our context is that of developing an approach to classify Landsat imagery to detailed land-cover category in persistently cloudy, complex tropical island landscapes for a multi-organizational project.

The problem with classifying cloud-filled or composite images is that the spectral signatures of each land-cover class will vary across such imagery by scene date. The resulting increase in signature variability increases spectral confusion among land-cover classes, limiting map accuracy and classification detail. The variability comes from interdate

Received 7 September 2006. Accepted 25 May 2007. Published on the *Canadian Journal of Remote Sensing* Web site at http://pubs.nrc-cnrc.gc.ca/cjrs on 23 August 2007.

E.H. Helmer.[1] USDA Forest Service, International Institute of Tropical Forestry (IITF), Jardín Botánico Sur, 1201 Calle Ceiba, Río Piedras, PR 00926, USA.

B. Ruefenacht. Red Castle Resources, USDA Forest Service Remote Sensing Applications Center, 2222 West and 2300 South, Salt Lake City, UT 84119, USA.

[1]Corresponding author. Present address: USDA Forest Service, IITF, c/o RMRS, 240 W Prospect Road, Fort Collins, CO 80526, USA (e-mail: ehelmer@fs.fed.us).

differences in atmospheric conditions, target illumination (sun–target–sensor geometry), sensor calibration, vegetation phenology, and soil moisture. These sources of interdate variability all change *image response function* (Schott, 1997). Corrections for atmospheric and illumination differences reduce interdate image differences, but they do not adjust differences in vegetation phenology. Image normalization also reduces interdate differences, but linear methods do not normalize nonlinear changes in vegetation phenology (Du et al., 2002). For imagery with coarse temporal resolution, years might pass before image acquisitions include scenes with both similar vegetation phenology and clear sky where clouds obscure land in an initial, reference scene. Because of these interdate image differences, classifications of image mosaics or composites often use machine learning classifiers, like neural networks or decision trees. These classifiers help accommodate data noise or spectrally heterogeneous classes (Friedl and Brodley, 1997; Seto and Liu, 2003). Two potential problems remain, however, when classifying such imagery. First, machine learning classifiers require that training data reflect class spectral variability. Training data from a limited number of field measurements may not always meet this requirement. Moreover, the limits to which machine learning classifiers can accommodate spectral heterogeneity have received little study. Second, interdate image differences, whether from phenological differences or residual errors after atmospheric correction, are often still visible in mosaics or composites. When large, these differences may make visual interpretation of mosaic imagery inaccurate, negatively impacting classifications that depend on visual interpretation.

The first specific objective of this study is to compare three potential methods for normalizing interdate image differences when filling cloudy areas in Landsat images with imagery from other dates. These methods include linear histogram matching, linear regression, and regression trees. Regression trees can form complex nonlinear relationships (Huang and Townshend, 2003). By forming regression models at the terminal nodes of decision trees, they have the potential to partition the relationships between spectral bands in differently dated scenes into sets of relationships for each band. Consequently, we hypothesize that they will more closely match vegetation phenology in images from different dates. Our second objective is to learn whether normalizing images to fill cloudy areas in reference scenes, which will presumably enhance visual interpretation, results in imagery that permits accurate mapping of land cover and detailed forest formations in a complex tropical landscape. We hypothesize that classifying cloud-filled imagery with decision trees will be accurate as long as (*i*) training data fully represent class variability, and (*ii*) ancillary geographical data are included with spectral bands in classifying the imagery. We also expect that having at least two cloud-filled images, each from a different season, will improve the classification.

Although many tropical islands are small, this study addresses Landsat imagery because it is nevertheless relevant for these landscapes. First, Landsat imagery has short-wave infrared bands that can be critical to mapping woody vegetation attributes in mountainous tropical landscapes (Helmer et al., 2000). Second, time series of imagery are sometimes critical to mapping certain land-cover and woody vegetation attributes and land-cover change. Time series of very fine spatial resolution imagery can become prohibitively expensive. Third, using Landsat imagery for mapping allows us to use satellite imagery with very fine spatial resolution for accuracy assessment. Lastly, not all complex tropical landscapes are small in area. Using very fine spatial resolution satellite imagery for land-cover and forest formation mapping over larger complex areas may be impractical for some time.

Importance of remote sensing methods focused on complex tropical landscapes

This study focuses on complex tropical landscapes, which we define as those that extend over several different forest formations (physiognomic and environmental) or forest types (compositional). In many tropical regions, for example, landscapes extend from lowland dry or moist areas to mountainous wet areas over short distances. The distances are particularly short on tropical islands. These landscapes compound the challenge of assembling imagery and classifying vegetation types over areas where clouds are persistent (Helmer et al., 2002). Over areas smaller than that of one Landsat scene, woody vegetation formations can change from drought deciduous woodlands and shrublands to semideciduous, evergreen, and cloud forests. Forest deciduousness timing, intensity, and spatial distribution can also vary from year to year. Moreover, tree species assemblages and forest formations do not necessarily vary continuously with climate, elevation, or geology. One reason is that disturbance also affects their distribution. In dry and dry–moist zones, for example, more recently disturbed tropical forest stands are more deciduous than nearby older forest (Condit et al., 2000; Arroyo Mora et al., 2005). At higher elevations, elfin cloud forest tree species may colonize disturbed lands at lower elevations instead of the species that compose adjacent older, taller cloud forest (Kappelle et al., 1995). Where biophysical factors do not predict the distributions of different forest formations because of disturbance, forest spectral response in satellite imagery may become more important when mapping forest types. Because forest types can vary similarly with elevation and disturbance in temperate landscapes, work in complex tropical landscapes has broader relevance.

Overcoming the remote sensing challenges that complex tropical landscapes pose is important for biodiversity conservation. Most gaps in the global network of protected areas are montane or insular regions in the tropics (Rodrigues et al., 2004). Conservation planning usually starts with mapping the distributions of different habitats, like different forest formations, by classifying Landsat satellite imagery (Scott et al., 1993). Although islands can be small, seven of 25 global "biodiversity hotspots" that Myers et al. (2000) identified are tropical islands or island chains. Nearly half of all these

hotspots include complex tropical landscapes with mountains. Most tropical islands are biodiversity hotspots because high species endemism combines with extensive habitat loss. The Caribbean region, the location of this study, is one of the "hottest hotspots". About 11.3% of the original primary Caribbean vegetation contains 2.3% and 2.9% of the world's endemic plants and vertebrates, respectively (Myers et al., 2000). Lowland and coastal ecosystems are relatively unprotected. At the same time, lowland ecosystems are subject to intense land development pressure (Lugo et al., 1981; Helmer, 2004). Remote sensing research can bias toward temperate or topographically simple continental sites (Castro et al., 2003). The problem of radiometric differences between satellite images captured on different dates exemplifies this tendency. Related work on simpler landscapes may not solve the problems specific to complex ones.

Background

Merging Landsat image dates across space

One solution to persistent cloud cover is compositing imagery with algorithms that select and merge across space pixels that are most likely to be cloud free. This solution is common with high temporal resolution imagery; image dates in these composites typically span short time periods of 5–32 days (Holben, 1986; Cihlar et al., 1996). The short time periods minimize phenological differences between input scenes. The daily or bi-daily image acquisition that consistent phenology requires for a composite is only widely available, however, for imagery with high temporal resolution (but coarser spatial resolution). The interdate differences across an image composite can also still complicate a classification. A few studies classify merges across space of images with coarser temporal resolution that span longer time periods. They address large image mosaics of Landsat imagery that join scene footprints after histogram matching (Homer et al., 1997) or atmospheric correction (Pax-Lenney et al., 2001). As with image composites, machine learning algorithms can accurately classify large mosaics or time series of mosaics to map land-cover (Vogelmann et al., 2001; Homer et al., 2004) or forest-cover change (Woodcock et al., 2001). Another example of filling gaps in imagery comes from filling scan gaps in *SLC-off* Landsat enhanced thematic mapper plus (ETM+) scenes, which are scenes dated after the scan line corrector (SLC) failed on the Landsat-7 instrument (Howard and Lacasse, 2004). Localized linear histogram matching normalizes one or more *subject* scene dates of thematic mapper (TM) or ETM+ imagery, which are *SLC-on* scenes, to fill scan gaps in a *reference* scene, which in this case is the SLC-off scene. The method predicts the digital number (DN) of each scan-gap pixel from the mean DN of a 17 pixel window in an SLC-on scene that is centered on the gap pixel. It corrects for differences in DN range and bias over the window from pixels that the two scenes have in common. The nonzero pixels of the SLC-off scene are present on the edges of the 17 pixel window (Scaramuzza et al., 2004). In applying a separate local histogram match for each pixel, the procedure is a type of nonlinear normalization when its image-wide effect is considered. Unlike cloud elements, though, the scan gaps are narrow and have fairly regular spacing. Although effective, the current method for producing gap-filled ETM+ imagery requires fairly clear imagery, in part because cloud spatial patterns are irregular. Because of that cloud pattern irregularity, this study only compares methods that use all overlapping, mutually clear pixels to estimate normalization models that may reduce interdate differences.

Only one related study considers imagery with coarser temporal resolution over persistently cloudy tropical landscapes. That recent work (Helmer and Ruefenacht, 2005) uses regression trees, which are nonlinear, to reduce interdate image differences between Landsat scenes. It then mosaics the cloud-free parts of images across space. It shows that two temporal end points of such cloud-filled imagery can support accurate urban change detection with a simple linear classifier. That study only estimates average error in the regression tree matching by band across the entire mosaic. It does not analyze errors by land-cover class, nor does it compare the regression tree approach with other approaches, as in this study. Moreover, it addresses only non-subtle change detection. In contrast, this study addresses classification to detailed forest formation.

Relative normalization of imagery

Previous work uses various methods to reduce radiometric differences between differently dated scenes. These methods may also permit subject image data to more seamlessly fill cloud gaps in reference scenes. The methods include absolute atmospheric correction and relative radiometric normalization. Atmospheric correction corrects for interdate differences in atmospheric optical depth, commonly converting image DNs to reflectance at the Earth surface. Relative radiometric normalization calibrates images to each other via the relationships between pixels from differently dated scenes. Usually relative radiometric normalization applies one normalization model across the entire scene for each date and for each band. Most previous work on relative radiometric normalization focuses on normalizing scenes in a time series to each other or to one reference scene in a way that avoids obscuring change detection (Schott et al., 1988; Vogelmann, 1988; Olsson, 1993; Song et al., 2001; Du et al., 2002). Consequently, these applications do one or more of the following: (*i*) only estimate normalization models from pixels of "spectrally invariant" features like man-made structures; (*ii*) exclude pixels that have changed land cover; (*iii*) use only same-season imagery; (*iv*) or use only linear models.

The landscape in this study includes both deciduous and evergreen forest types and cultivated lands and pastures. Consequently, we make separate cloud-filled images for each of two stages in vegetation phenology, namely when deciduous forest trees are leaf-on and when they are leaf-off. Presumably, stacking images from different stages in vegetation phenology may help distinguish classes. Such imagery has been referred to

as *multiseason* imagery. It might merge, for example, the six optical bands from three Landsat ETM+ scenes to form an 18 band image. The 18 band image might include leaf-off, green-up, and peak growing season data. Previous work has shown that in temperate landscapes such multiseason satellite imagery can improve land-cover classifications (Lo et al., 1986; Schriever and Congalton, 1995; Wolter et al., 1995). Cloud-free imagery is most commonly available during the middle of each dry season, which is when drought deciduous vegetation may be greened up. However, the availability of cloud-free imagery over a year varies.

Methods

Overview

To compare potential methods for normalizing imagery that fills cloud gaps, we compared how closely the methods matched the spectral signatures of land-cover and forest formation classes in Landsat scenes from different dates. In other words, we analyzed the differences between reference-scene signatures and signatures for corresponding pixels in other scenes after normalization to the reference scene. We separately compared these differences by land-cover class. For example, we compare results for evergreen forest, drought deciduous forest, and agriculture. These three classes have different seasonal patterns in vegetation phenology. That is, they differ in the degree and direction of spectral change over time. In addition to comparing the mean absolute differences between reference and normalized imagery, we also present visual comparisons of cloud-filled images. Lastly, we compared the accuracies of classifying cloud-filled imagery developed with the different methods.

Study area

The country of St. Kitts and Nevis and the island of St. Eustatius, which is one of the Netherlands Antilles, are Leeward Islands in the Caribbean Lesser Antilles (**Figure 1**). They occupy a portion of Worldwide Reference System (WRS) path 002, row 048. Together these three islands occupy about 290 km^2. These islands include many of the land-cover types and forest formations found on other islands of the Lesser Antilles, yet they are relatively small. These two attributes make the islands a good area for testing methods to map land cover in the Lesser Antilles. Consequently, we used these islands for developing methods in support of a multi-organizational project. The project goal is to classify Landsat images to land cover and detailed forest formation for most islands in the region.

Each of the three islands has one or more mountains of volcanic origin, and elevations range from sea level to about 1156 m. Interactions between this topography and trade winds from the northeast have yielded complex natural vegetation. Woody vegetation formations range from subtropical drought deciduous and semideciduous woodlands and forest, including mosaics of xeric coastal formations that have evergreen, deciduous, or mixed forest and shrublands, to more humid broadleaf subtropical evergreen forests including cloud forests. Drought deciduous woody vegetation includes woody formations in which at least 75% of woody canopy trees are deciduous. About 25%–75% of canopy trees are deciduous in semideciduous forest, and at least 75% of canopy trees are evergreen in seasonal evergreen and evergreen forest (sensu FGDC, 1997). Drought deciduous woodlands include lands with >25% canopy cover of drought deciduous shrubs or trees, which are often leguminous and thorny, and an understory of grasses and forbs that fire or grazing maintain. If these disturbances cease, drought deciduous woodlands succeed to drought deciduous forest, and they may eventually succeed to semideciduous or mixed coastal forests.

Relative normalization for filling clouds in base scenes: overview

We selected Landsat-7 ETM+ image dates based on the goal of developing two final cloud-filled images for the study area. Each of the two cloud-filled images would have drought deciduous woody vegetation in a different stage of phenology. Another goal was to minimize the number of scenes that would compose each cloud-filled image. Presumably, a smaller number of image dates should limit the spectral variability of each land-cover class. Using these criteria, the image dates that we selected were the four clearest scenes available at the time of the study that together would (*i*) yield the most cloud-free area over the islands, (*ii*) include some clear areas in common (for estimating the normalization models), and (*iii*) form the two cloud-filled images. Because phenology varies from year

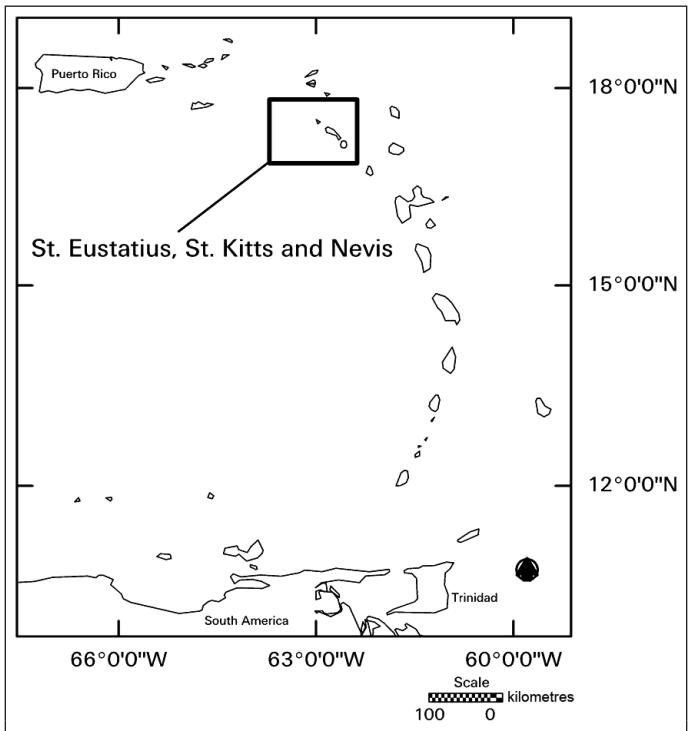

Figure 1. Location of study area in the West Indies.

to year, as mentioned previously, the overlay order of the various image dates in each cloud-filled image was not based on date alone. The four scenes are dated 12 December 1999, 5 September 2000, 11 September 2002, and 2 February 2003 (**Table 1**). Georeferencing for the two earliest scenes included terrain correction that corrected parallax error from local topographic relief with a digital elevation model (Level 1T, http://landsat7.usgs.gov/productinfo.html). They coregistered to within one 30 m pixel, which additional coregistration could not improve. We manually collected ground-control points to coregister the two later scenes to the least cloudy, earliest scene.

Drought deciduous woody vegetation is mostly in a leaf-on stage in the earliest scene, and lowland pasture is greened up except in the driest areas (leaf-on scene). This fairly clear scene served as the reference scene for the cloud-filled image with leaf-on drought deciduous woody vegetation. The second scene was the primary scene that filled cloudy areas for the leaf-on cloud-filled image. Its overall phenology was closest to that of the reference leaf-on scene. Drought deciduous forest is leaf-on in this second scene, though some xeric coastal forest and drought deciduous woodlands are leaf-off (*partially leaf-on* scene). For the cloud-filled image with deciduous woody vegetation in a leaf-off stage, the most cloud-free of the two leaf-off scenes was used as the reference scene. Drought deciduous forests, woodlands, and shrublands are all leaf-off in these latter scenes, and pasture grass is senescent (except in humid zones at high elevations) (**Table 1**).

A few cloud gaps remained after mosaicking cloud-free parts from the initial reference and subject scenes for each of the two cloud-filled images (one leaf-on and one leaf-off). Most of the remaining cloud gaps were filled with data from the other two image dates that were normalized to the appropriate reference scene in the overlay order shown in **Table 1**. Normalizing three image dates to each of the two reference scenes resulted in a total of six subject-to-reference scene pairs. For example, in the case of the leaf-on scene being the reference scene, each of the three later scenes underwent normalization to that reference scene. Each subject-to-reference scene pair underwent each of three different normalization methods that we detail later: regression trees, linear regression, and histogram matching. The average absolute differences between reference scene pixels and normalized or atmospherically corrected subject scene pixels were then compared for different vegetation types. We also visually compared the mosaics with each other and with mosaics assembled from images that underwent simple image-based atmospheric correction. We rescaled the estimated reflectance values to an eight-bit scale comparable to the other DN ranges.

The relative normalization approaches that we compare assume that each reference image and subject image have some cloudless areas in common. Each approach separately models the relationship between one band in a reference scene and one or more bands in a subject scene. These models allow subject scene data to predict reference scene data from all pair-specific, mutually clear pixels. To find overlapping cloud-free pixels, we first made cloud and cloud shadow masks (cloud – shadow masks) and then masked most ocean water. The union of the cloud – shadow masks from each reference–subject scene pair then masked both scenes, which revealed where both scenes are cloud free. After estimating the normalization models, we applied them to all cloud-free parts of each subject scene to compare the signatures of mutually cloud-free areas. To construct cloud-filled imagery, normalized subject scene data fill cloudy areas in reference scenes. The subject scene that was phenologically most similar to the reference scene fills cloudy areas first, as mentioned earlier. Data from other normalized subject scenes fill remaining cloudy areas (**Table 1**). Comparisons of class spectral signatures focus on the two clearest reference scenes and their differences with normalized subject scenes. To gauge the effect of cloud cover amount, however, comparisons of classification accuracy include mosaic images that used the two cloudiest scenes as reference scenes.

Table 1. Vegetation phenology of scene dates for Landsat ETM+ scenes in this study and overlay order of subject image data used to fill clouds in cloud-filled images constructed for a given reference scene date.

Phenology of drought deciduous woody vegetation[a]	Phenology of pasture or grass in lowlands	Reference scene date (top)	Overlay order of subject scenes (by scene date) for multipart image[b]	Cloud-obscured land cover (%)
Reference scenes for cloud-filled images with fewest clouds in each reference scene				
Leaf-on	Mostly green	12 Dec. 1999[c]	5 Sept. 2000 – 2 Feb. 2003 – 11 Sept. 2002	9.1
Leaf-off	Senescent	11 Sept. 2002[d]	2 Feb. 2003 – 5 Sept. 2000 – 12 Dec. 1999	20.7
Reference scenes for additional cloud-filled images				
Partially leaf-on[e]	Senescent	5 Sept. 2000	12 Dec. 1999 – 11 Sept. 2002 – 2 Feb. 2003	36.5
Leaf-off	Senescent	2 Feb. 2003	11 Sept. 2002 – 12 Dec. 1999 – 5 Sept. 2000	50.7

[a]Drought deciduous woody vegetation includes drought deciduous woodlands, drought deciduous forest, and those xeric coastal forest patches that are drought deciduous. The xeric coastal forests include various mixtures of sclerophyllous evergreen, deciduous, and succulent woody species, as well as grasses.

[b]The subject images are given in order of overlay as second – third – fourth (bottom).

[c]Reference–subject scene pairs in mosaic with 12 Dec. 1999 scene as reference scene: 12 Dec. 1999 – 5 Sept. 2000, 12 Dec. 1999 – 2 Feb. 2003, and 12 Dec. 1999 – 11 Sept. 2002.

[d]Reference–subject scene pairs in mosaic with 11 Sept. 2002 scene as reference scene: 11 Sept. 2002 – 2 Feb. 2003, 11 Sept. 2002 – 5 Sept. 2000, and 11 Sept. 2002 – 12 Dec. 1999.

[e]Drought deciduous forest is generally greened up, but woodlands and some xeric coastal forest patches are not.

Regression tree models

As mentioned previously, the normalizations model the relationships between each band in a reference scene and one or more bands in a subject scene. The resulting models then predict new values of subject scene pixels that are calibrated to the reference scene. For regression tree normalization, earlier work indicated that models with many predictor bands had less error (Helmer and Ruefenacht, 2005). Consequently, the regression tree model for each band and reference–subject scene pair has the following general form:

$$y_{refi} = f(x_{subj1}, x_{subj2}, x_{subj3}, x_{subj4}, x_{subj5}, x_{subj7}) \quad (1)$$

where y_{refi} is the DN of a pixel in the reference scene for the ith band to be predicted; and x_{subj1}–x_{subj5} and x_{subj7} are the DNs of ETM+ bands 1–5 and 7, respectively, of the corresponding pixel in the subject scene. After estimating a model from reference and subject scene areas that are mutually clear, the model can estimate new DNs for locations with clear subject scene data but obscured or absent reference scene data as follows:

$$y_{matchi} = f(x_{subj1}, x_{subj2}, x_{subj3}, x_{subj4}, x_{subj5}, x_{subj7}) \quad (2)$$

In the procedures, overlapping clear areas first supply bands for six new images, one for each optical band of the reference scene (Helmer and Ruefenacht, 2005). Each of the six images contains the dependent and independent variables for a regression tree model that predicts one of the reference scene bands, as in Equation (1). For example, the image for the model to predict band 1 in the reference scene will have the following seven bands: y_{ref1}, x_{subj1}, x_{subj2}, x_{subj3}, x_{subj4}, x_{subj5}, and x_{subj7}, where the notation is the same as that for Equation (1). The procedure uses an ERDAS Imagine interface to export each of these seven-band images to an ASCII file formatted for the regression tree software Cubist (www.rulequest.com). In formatting the data for Cubist, each row represents one pixel location and contains comma-separated brightness values for the band that is the dependent variable and the six bands that are the independent variables. The next step is to develop regression tree models for the bands within the regression tree software. We then apply each regression tree model to all clear subject scene areas. Integrating public domain code (www.rulequest.com) into an Imagine program, with the Imagine C Developer's Toolkit (Leica Geosystems GIS & Mapping, LLC, 2003), enabled Imagine to interpret and apply the six regression tree models that predicted new DNs for each pixel.

Linear models

We developed linear regressions with all mutually clear image pixels, rather than excluding from that pixel set areas of marked change as in some previous work (Song et al., 2001), because using all image pixels is more adaptable to fully automated cloud-filled image production. The linear relative normalization models consequently used the same sets of mutually cloud-free pixels as the regression tree models, but they had the following general form:

$$y_{refi} = f(x_{subji}) \quad (3)$$

where y_{refi} and x_{subji} are brightness values for the ith band from pixels in the reference and subject scenes, respectively; and the function $f(x_{subji})$ is linear. Equation (3) then estimates each radiometrically normalized pixel, y_{matchi}:

$$y_{matchi} = f(x_{subji}) \quad (4)$$

With linear regression models, Olsson (1993) found that using several subject image bands to predict each reference image band only slightly improved calibration models, and most previous applications limit the independent variables in linear normalization models to one corresponding subject image band. Some applications replace x_{subji} in Equations (1) and (2) with principal component axes of (*i*) multiple bands from the subject scene (Olsson, 1993) or (*ii*) a single band from multiple dates (Du et al., 2002).

Histogram matching

Histogram matching determines a look-up table for each image band that cross-references an output, matched DN for each subject image input DN. The matching equates the cumulative distribution functions of the reference and subject image histograms, which reveals a reference image DN to assign to subject image pixels on their conversion back to a frequency distribution (Richards, 1993). For histogram matching, we used an ERDAS Macro Language program, Image Match (Helmer and Ruefenacht, 2005). This program first matches only the histograms from overlapping parts of the subject and reference images using the ERDAS RASTERMATCH function (Leica Geosystems GIS & Mapping, LLC, 2003), which is a linear histogram match. It then uses the resulting output look-up table to assign a histogram-matched DN to each DN in the entire subject image. The advantage of Image Match is that it yields more closely matched histograms. By initially matching only image overlap areas, which by definition have equal pixel numbers and cover similar terrain, the histogram matching algorithm outputs look-up tables that have avoided scaling errors and minimize errors from differences in histogram shape.

Atmospheric correction

Atmospheric correction provides here a baseline when visually comparing the different image normalization approaches and when comparing overall classification accuracies because differences in vegetation phenology should not influence the correction. We applied a simple image-based approach to atmospheric correction, called dark object subtraction (DOS) (Chavez, 1996), to all image dates. "Minimum" DN values were selected from the histogram for each band at the point just before the histogram rises sharply,

and they had at least 100 pixels less than or equal to their value. A 100 pixel minimum was appropriate because the study area represented only a portion of a Landsat scene. Digital numbers that had 500–1000 pixels less than their value, which may be more appropriate for full Landsat scenes, were well above the lower inflections of the histograms for each band. As mentioned, we rescaled the resulting reflectance values to an eight-bit scale to make them visually comparable to the other cloud-filled images.

Reference data

We collected three sets of reference data locations for analyses. First, classification-training data identified land cover or vegetation formation of 25 to >100 multipixel patches in the satellite imagery. We distributed the training data throughout the extent of each class, including in cloud-filled areas. Field-based training data collection occurred from 8 to 21 January 2003. It relied on simultaneously observing land cover and forest formation both in satellite imagery and in the field by integrating a global positioning system (GPS) receiver with a laptop computer (with a daylight-viewable image display) running the ERDAS Imagine GPS tool (Leica Geosystems GIS & Mapping, LLC, 2003). We supplemented these field data by visually interpreting several pan-sharpened, false color IKONOS images dated around the year 2000. The resulting dataset trained a See5 (www.rulequest.com) decision tree model that classified a stack of raster data, yielding a land cover – forest formation map for the islands that we manually edited to make as accurate as possible. The raster data included the two cloud-filled images built from the two clearest reference scenes, one leaf-on and one leaf-off, and several layers of ancillary data (see next section).

The second reference dataset came from a stratified random sample of the map mentioned in the previous paragraph of 500 pixel locations per land-cover or forest formation class. This dataset supplied data for comparing normalized subject image signatures to corresponding reference image signatures. Third, a similar stratified random sample of about 50 pixels per class provided data for estimating the accuracies of classifying various stacks of cloud-filled images. We identified the actual land cover or forest formation of these pixels by visually interpreting the IKONOS imagery. For comparing the accuracies of classifying the different cloud-filled images, no manual editing was applied.

Evaluation of normalization methods

Three steps assessed the effects of the different normalization methods for viewing and classifying cloud-filled imagery. The first step compared reference image pixels and corresponding normalized or non-normalized subject image pixels. Second, we present visual comparisons of cloud-filled images developed using different normalization methods and DOS atmospheric correction. Third, we classify the cloud-filled images developed using the different methods and compare the resulting classification accuracies.

The reference dataset that randomly selected 500 pixels per class provided data for comparing mean differences between reference images and normalized subject images. For each reference–subject image pair, the actual number of observations is less than 500 because the analysis dropped observations that were not mutually clear. We performed one analysis of variance (ANOVA) with multiple comparisons of means (MCM) for each unique combination of ETM+ band or index, land-cover class, and reference–subject image pair. The explanatory variable of "method" for each ANOVA had four possible conditions: no normalization or correction, linear regression normalization, image match histogram matching, or regression tree normalization. The ETM+ bands–indices included bands 1–5 and 7, the normalized difference vegetation index (NDVI), and the band 4:5 ratio. The NDVI gauges vegetation greenness, and the band 4:5 ratio is sensitive to forest structure and successional stage (Fiorella and Ripple, 1993), including in tropical landscapes (Helmer et al., 2000). The eight bands, when multiplied by 10 classes and by six reference–subject image pairs, resulted in 480 ANOVAs for comparing mean absolute differences.

To learn how the normalization method affected classification accuracy, we used the training data locations to classify cloud-filled images. This step classified single cloud-filled images or stacks of cloud-filled images with their own signatures. As mentioned previously, the training data represented the extent of each class, including in cloud-filled areas. The data trained a See5 decision tree classification model based on Landsat bands 1–5 and 7 from one or more cloud-filled images and the ancillary raster data. The ancillary data included distance to primary road, distance to coast, distance to ravine, a discrete variable that distinguished each island, and topographic variables from Shuttle Radar Topography mission data (Farr and Kobrick, 2000). Topographic variables included elevation, slope, slope position, aspect, average aspect over 510 m (to yield topographic facet relative to the direction of trade winds), and topographic shading based on the sun elevation and azimuth of each reference image (Leica Geosystems GIS & Mapping, LLC, 2003). The accuracy assessment dataset supported estimates of Kappa coefficients of class agreement (Cohen, 1960). Preliminary work showed that classifying the cloud-filled images to detailed woody formation was inaccurate without ancillary data, yielding Kappa coefficients <0.6.

Results

Radiometric normalization methods significantly reduce the mean absolute differences between reference scene pixels and corresponding pixels from subject scenes that may be used to fill cloud gaps, as expected (**Figure 2**, see description later in this section). In addition, regression tree normalization most consistently yields the smallest significant differences between reference and subject images in ETM+ bands 1–5 and 7, NDVI, and the band 4 to band 5 ratio. These conclusions are based on the MCM tests (**Figure 2**) and the graphs of selected

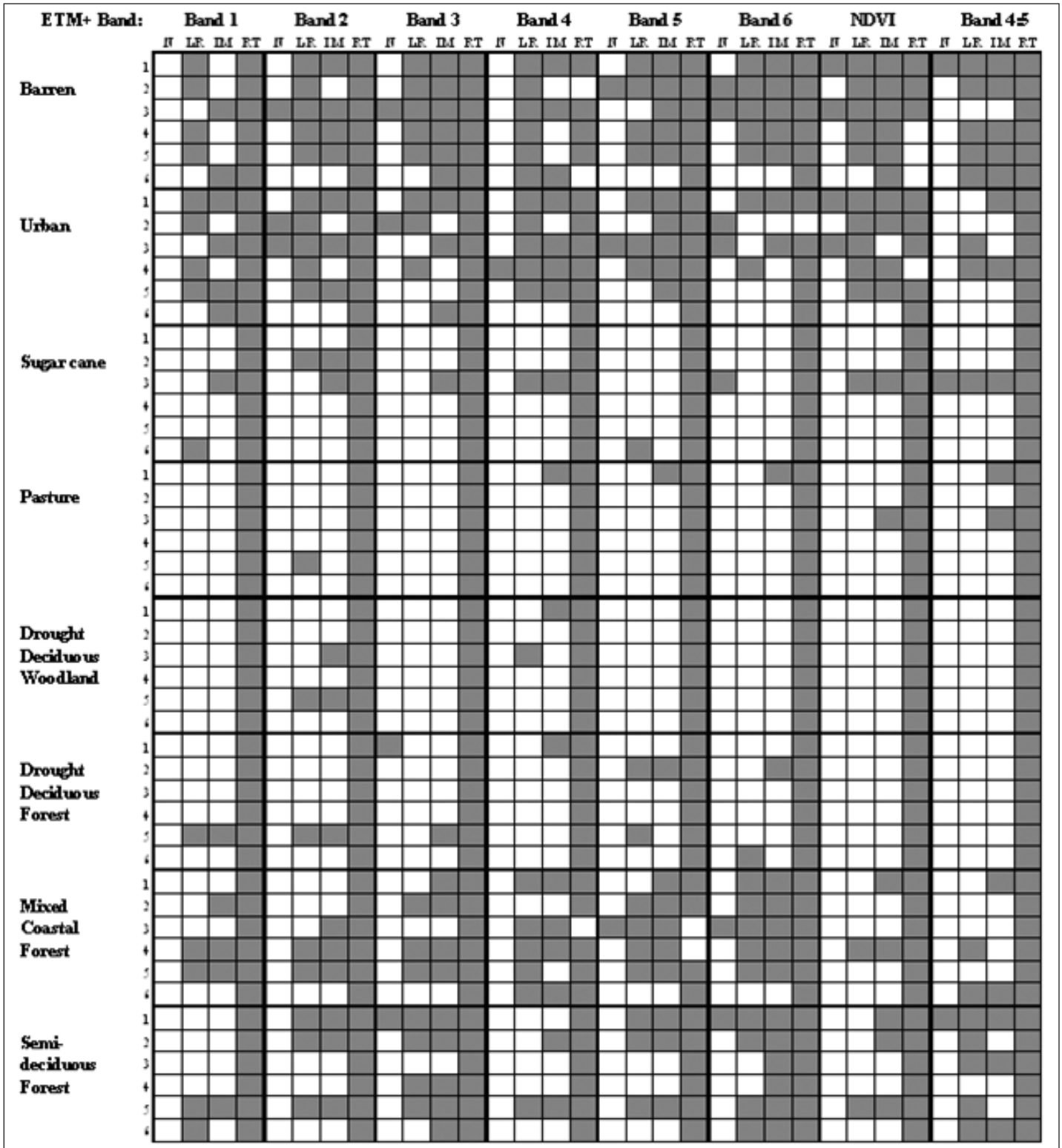

Figure 2. Results from 480 ANOVAs to compare normalization results. Each square outlined by a bold line shows results for one band and one class for each of six separate multiple comparisons of means (MCMs). The six MCMs result from the six reference–subject image pairs that fill clouds in the two cloud-filled images in **Figures 4** and **5** (three pairs for the leaf-on imagery in **Figure 4** and three pairs for the leaf-off imagery in **Figure 5**). The four small columns in each bold-outlined square are the four explanatory variable conditions: no normalization or correction (N), linear regression normalization (LR), image match histogram matching (IM), or regression tree normalization (RT). Shaded squares indicate which conditions yielded the smallest significantly different group of means for each MCM (for the particular band, class, and reference–subject image pair). Reference–subject scene pairs 1–6 are as follows: 1, 11 Sept. 2002 – 2 Feb. 2003; 2, 11 Sept. 2002 – 5 Sept. 2000; 3, 11 Sept. 2002 – 12 Dec. 1999; 4, 12 Dec. 1999 – 2 Feb. 2003; 5, 12 Dec. 1999 – 5 Sept. 2000; 6, 12 Dec. 1999 – 11 Sept. 2002.

differences (**Figure 3**). Visually evaluating the different cloud-filled images is consistent with the MCM results (**Figures 4, 5**). When subject image data that fill cloud gaps in a reference image undergo normalization with regression trees, the resulting cloud-filled imagery is more seamless than when images undergo histogram matching or linear regression normalization. This trend is most apparent for vegetation that undergoes marked phenological change (including sugar cane, pasture, drought deciduous woodland, and drought deciduous forest) than for evergreen or mixed forest formations. Cloud-filled imagery made with all the normalization approaches, especially regression tree normalization, is also more seamless than when reference and subject scenes both undergo DOS atmospheric correction, including for ETM+ bands 1, 2, and 7, in addition to bands 3, 4, and 5 displayed in **Figures 4** and **5**. Normalizing image data with regression trees helps to match the phenology of vegetation in scenes that have fairly different phenology (**Figure 6**).

Although **Figure 2** is detailed, its shading strategy is meant to convey the overall results of the 480 MCMs without the need to evaluate each individual result. Each square outlined by a bold line in **Figure 2** shows results for each of the six separate MCMs for one band and one class. The six MCMs are for the six different reference–subject image pairs, which are labeled as small rows one through six. The shaded cells in each small row indicate which methods are in the smallest significantly

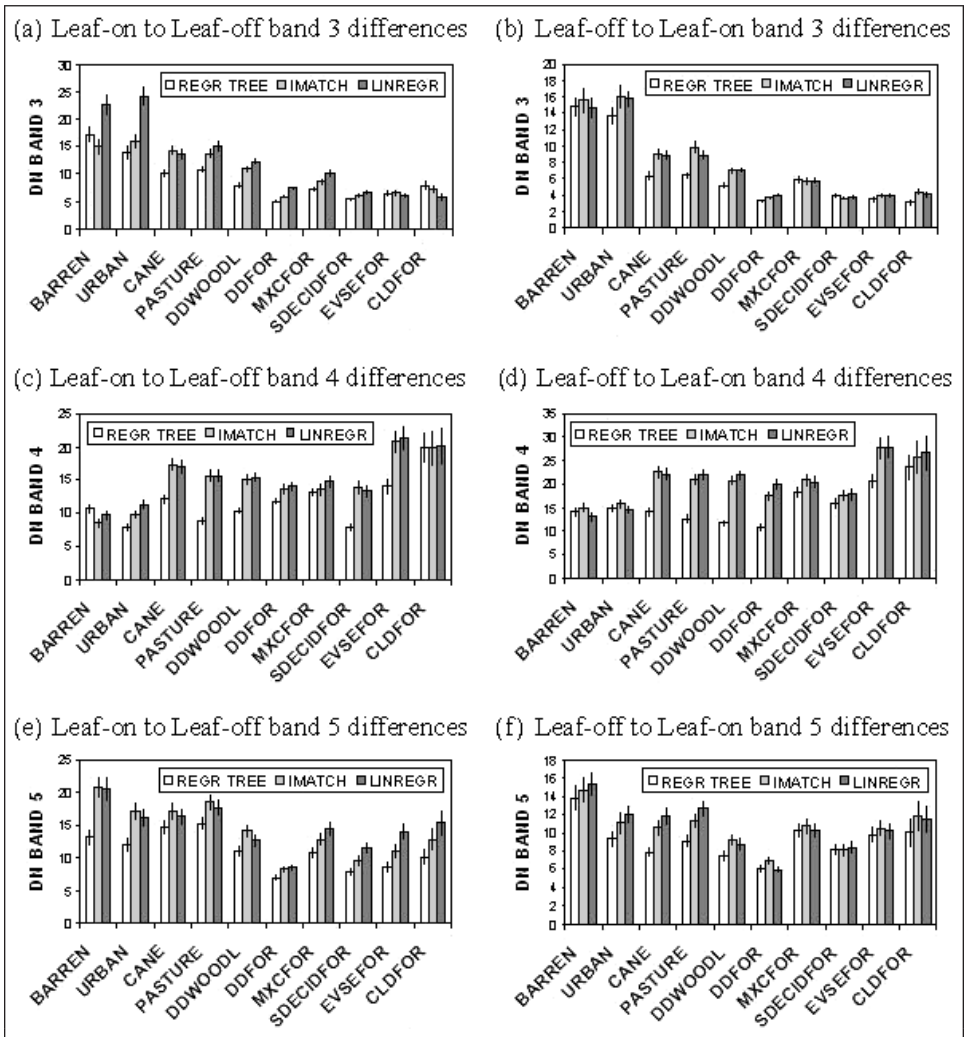

Figure 3. Mean absolute differences and 95% confidence intervals for Landsat ETM+ band 3 (a and b), band 4 (c and d), and band 5 (e and f) for two reference–subject image pairs, including 12 Dec. 1999 – 11 Sept. 2002 (leaf-on to leaf-off) and 11 Sept. 2002 – 12 Dec. 1999 (leaf-off to leaf-on) for the following classes: barren (BARREN), high-density urban (URBAN), sugar cane (CANE), pasture and other grass (PASTURE), drought deciduous woodland (DDWOODL), drought deciduous forest (DDFOR), mixed coastal forest (MXCFOR), semideciduous forest (SDECIDFOR), evergreen and seasonal evergreen forest (EVSEFOR), and cloud forest (CLDFOR). Mean differences after the following treatments are shown: regression tree normalization (REGR TREE), image match histogram matching (IMATCH), and linear regression normalization (LINREGR).

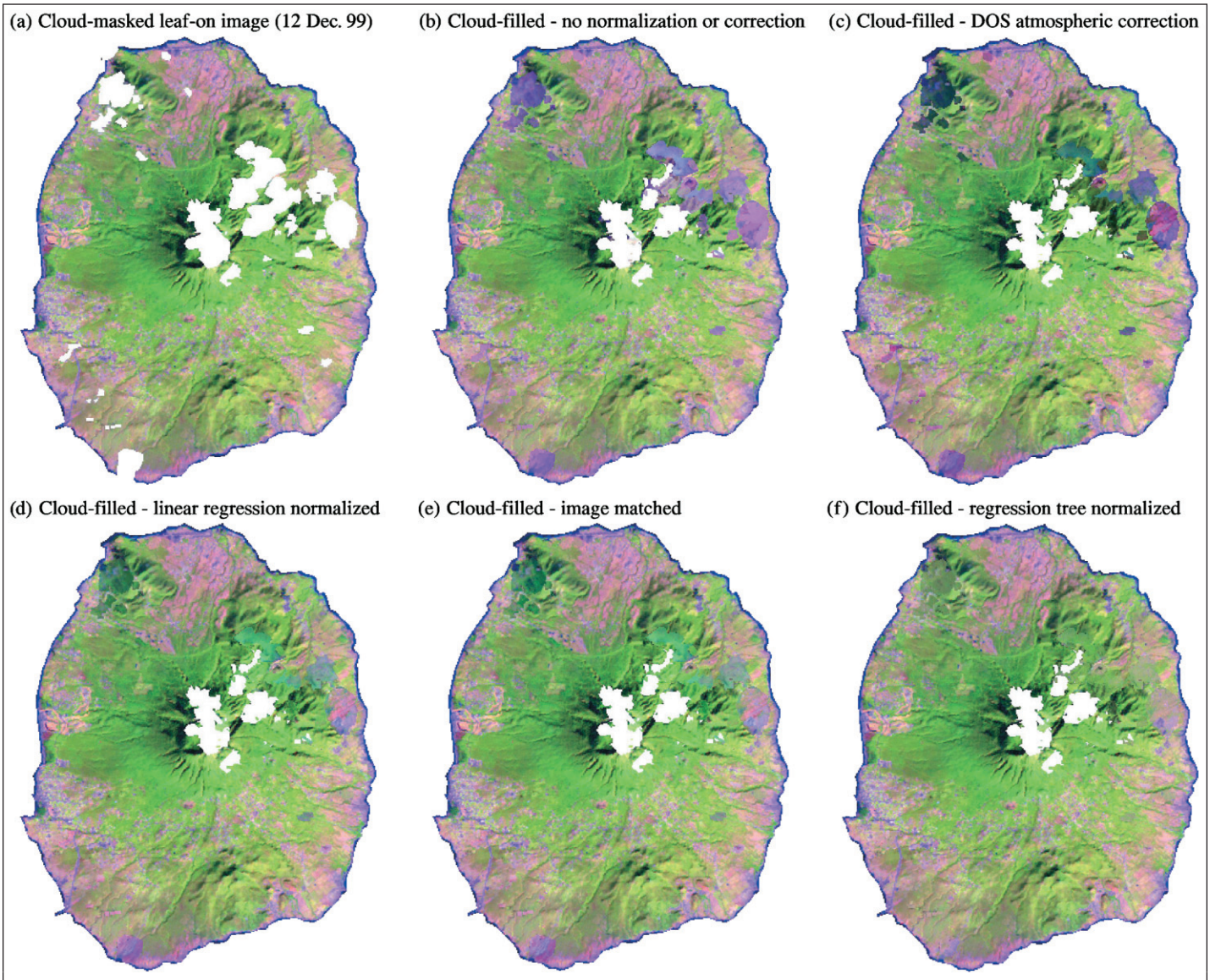

Figure 4. The island of Nevis, with Landsat ETM+ bands 3, 4, and 5 displayed in red, green, and blue, respectively, in the clearest leaf-on scene (12 Dec. 1999) with (a) clouds masked and (b–f) clouds filled by other image data that were subject to no normalization or atmospheric correction (b), atmospheric correction with dark object subtraction (c), normalization with linear regression (d), normalization with image match linear histogram matching (e), and regression tree normalization (f). **Table 1** gives the overlay order (top to bottom) of image dates. Color tables are identical for all image data.

different group of mean absolute differences between reference and subject images. For example, all of the six MCM results for ETM+ band 1, class barren, are in the upper left bold-outlined square of **Figure 2**. The four small columns in each bold-outlined square are the four explanatory variable conditions, namely no normalization or correction (N), linear regression normalization (LR), image match histogram matching (IM), or regression tree normalization (RT). The MCM result for the first image pair (row 1) is that two of the normalization approaches (LR and RT) yielded reference–subject image differences that are significantly smaller than no normalization or image match histogram matching. However, the mean differences from the regression tree and linear regression normalization methods are not significantly different from each other. Consequently, the two small squares in small row one under those methods are shaded. The pattern of shading in **Figure 2** shows that the RT columns, in the far right of each bold-outlined column, are all shaded except in a few cases. The RT columns are most often the only ones shaded for agricultural, grassland, and deciduous woody classes, which are the most phenologically variable.

Another way of viewing this information is by graphing mean absolute differences between subject and reference images as in **Figure 3**. The data in **Figure 3** are a subset of the comparisons in **Figure 2**, and they compare the same two reference–subject image pairs, in bands 3, 4, and 5, as compared in **Figure 6**. Again in **Figure 3**, the smallest differences between reference and subject images tend to be for

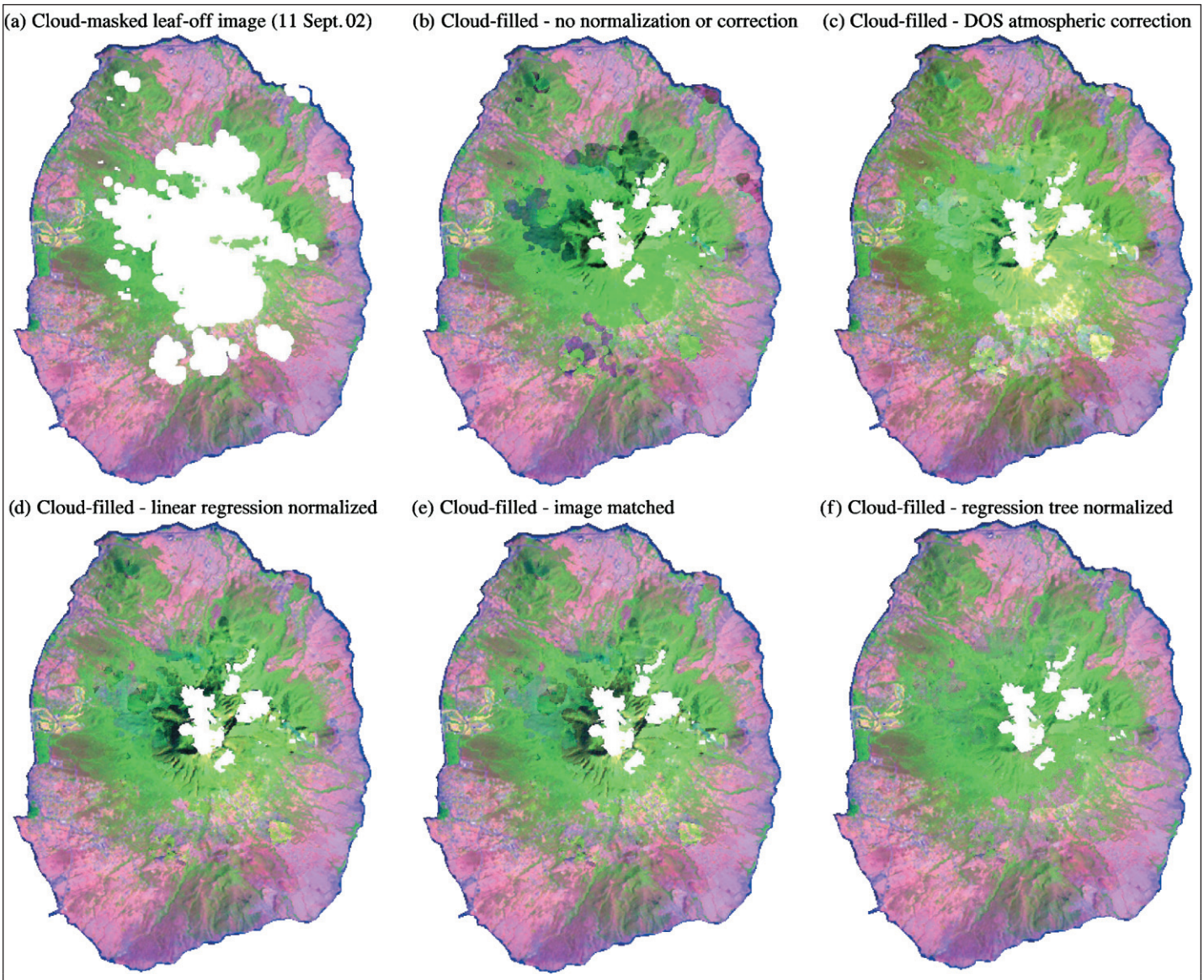

Figure 5. The island of Nevis, with Landsat ETM+ bands 3, 4, and 5 displayed in red, green, and blue, respectively, in the clearest leaf-off scene (11 Sept. 2002) with (a) clouds masked and (b–f) clouds filled by other image data that were subject to no normalization or atmospheric correction (b), atmospheric correction with dark object subtraction (c), normalization with linear regression (d), normalization with image match linear histogram matching (e), and regression tree normalization (f). **Table 1** gives the overlay order (top to bottom) of image dates. Color tables are identical for all image data.

subject images that underwent regression tree normalization, especially for the phenologically most variable classes.

The accuracies of classifying the cloud-filled images developed with various methods do not significantly differ, as gauged with the Kappa coefficient (**Table 2**). This result is at first counterintuitive. Apparently, if classifications use decision trees, ancillary data, and well-distributed training data that include cloud-filled pixels, the different approaches to normalizing the image data that fill cloud gaps in these landscapes do not significantly affect overall classification accuracy. Classifications of cloud-filled reference scenes with less initial cloud cover tend to be more accurate than if reference scenes are cloudier, but the differences between Kappa coefficients are not significant for this landscape (**Table 2**). These results are not likely to be artifacts of the accuracy assessment data because the maps resulting from the various classifications did not greatly differ.

Having cloud-filled images from two different stages of vegetation phenology (multiseason imagery) greatly enhances the ability to distinguish different types of deciduous forest when collecting training data. The multiseason imagery does not significantly improve overall classification accuracy in this particular landscape. However, per-class accuracies for selected image combinations and normalization methods (**Table 3**) suggest that the main difference between classifications that use one image date and those that use two dates is that producer's accuracy increases slightly with classifications that use two image dates for some

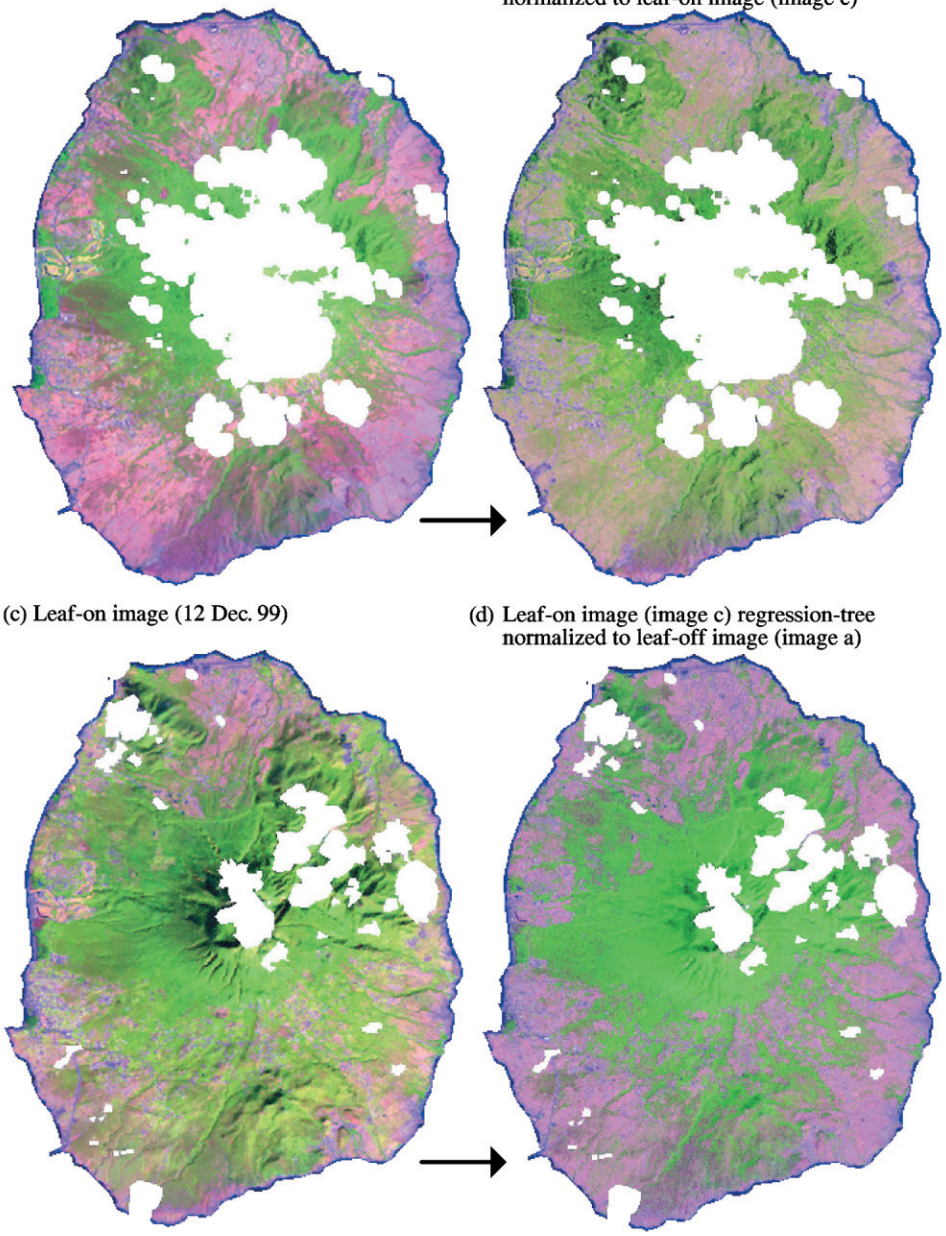

Figure 6. The island of Nevis with clouds masked in (a) the clearest leaf-off scene (11 Sept. 2002), (b) the leaf-off scene (image a) regression-tree normalized to the leaf-on scene (image c), (c) the clearest leaf-on scene (12 Dec. 1999), and (d) the leaf-on scene (image c) regression-tree normalized to the leaf-off scene (image a). All images displayed in Landsat ETM+ bands 3, 4, and 5 in red, green, and blue, respectively. The color tables for images (a) and (d) are identical, as are those for images (b) and (c).

phenologically variable classes: pasture, drought deciduous woodland, and semideciduous forest (though accuracy slightly decreases for some other classes). In addition, with two image dates, slight increases in per-class accuracy are more common for classifications that use normalized imagery than when images are not normalized. However, the differences in per-class accuracies vary by class, indicating that the differences in overall classification accuracy between normalization methods will depend on the landscape under study.

Discussion

When filling cloudy areas in Landsat imagery with imagery from other dates, normalizing subject to reference scene data reduces interdate image differences. Normalization results in more seamless cloud-filled imagery than simple DOS atmospheric correction of reference and subject scenes. This result holds for all ETM+ bands in this Caribbean island landscape. Regression tree normalization tends to reduce these

Table 2. Kappa coefficients of class agreement (±95% confidence intervals for Kappa) that resulted from classifying different cloud-filled images and ancillary data did not differ significantly from each other.

	Normalization or correction method applied to image data				
Reference image for the cloud-filled image(s) classified	Cubist regression tree normalization	Image match histogram matching	Linear regression normalization	DOS atmospheric correction	No normalization or correction
Sept. 2000	0.64±0.08	0.64±0.08	0.64±0.08	0.64±0.08	0.63±0.08
Feb. 2003	0.67±0.08	0.66±0.08	0.67±0.07	0.66±0.07	0.66±0.07
Feb. 2003, Sept. 2000	0.66±0.07	0.66±0.07	0.64±0.08	0.66±0.07	0.66±0.07
Sept. 2002 (clearest leaf-off)	0.66±0.07	0.69±0.07	0.66±0.08	0.68±0.08	0.65±0.08
Dec. 1999 (clearest leaf-on)	0.71±0.08	0.68±0.07	0.69±0.07	0.70±0.07	0.68±0.08
Sept. 2002, Dec. 1999 (two clearest)	0.72±0.07	0.71±0.07	0.70±0.07	0.71±0.07	0.68±0.07
Feb. 2003, Sept. 2000, Sept. 2002, Dec. 1999	0.70±0.07	0.72±0.07	0.69±0.08	0.70±0.07	0.69±0.07

Note: The result indicates that with well-distributed training data, decision tree classification, and ancillary data the various treatments applied to subject image data used to fill clouds in reference scenes may not always significantly affect classification accuracy.

differences more than linear regression or histogram matching because it more closely matches vegetation phenology. These results are visible in the relative seamlessness of cloud-filled imagery that displays, for example, ETM+ bands 3, 4, and 5. These results also support the hypothesis that regression tree normalization can result in more seamless cloud-filled imagery.

Cloud-filled Landsat imagery can apparently be classified to land cover and detailed forest formations in complex tropical landscapes with the methods in this study. This result supports our second hypothesis. The success of the classifications probably stems from three main factors: the well-distributed training data, the use of decision tree classification, and the ancillary data. First, the well-distributed classification training data reflected the spectral and spatial variability of each class, including the variability that stemmed from interdate image differences, because the training data included cloud-filled areas. If the training data represent most spectral conditions of a land-cover class, the decision trees can apparently resolve some of the spectral incongruities of cloud-filled imagery. Second, the training data reflected the spatial variability of each class. Topographic derivatives and other ancillary data help distinguish land cover and forest types in these landscapes. Ancillary data probably also helped to detect differences in forest deciduousness related to disturbance because topography and road proximity are also related to the spatial patterns of tropical forest disturbance and regrowth (Helmer, 2000). The fact that the spatial distributions of land cover and forest formations are related to topography and other ancillary data probably also explains why multiseason imagery did not improve classification accuracy as much as expected.

In this study, neither normalizing nor atmospherically correcting image data that filled cloud gaps significantly improved or degraded classification accuracy. This result can be surprising, because the different normalization methods usually reduced differences between reference scenes and the subject image data that filled cloud gaps. A possible explanation for this result is that decision trees can handle the spectral heterogeneity of cloud-filled imagery, and normalizing the imagery that fills cloud gaps increases spectral confusion for some classes, offsetting gains in accuracy for other classes. Evidence for this explanation lies in the structures of the decision trees, the variability in normalization effects on per-class accuracy, and the differences between classifications with ancillary data versus those without ancillary data. For a related study, we analyzed the decision trees that we ultimately used for mapping land-cover and forest formations for St. Kitts and Nevis and St. Eustatius, as well as for the Caribbean island of Grenada (E.H. Helmer et al., unpublished data). The variables that most commonly appear in the top nodes of the decision trees are spectral, or they are ancillary variables that affect the spectral bands, like image topographic shading based on image sun–target–sensor geometry. This outcome suggests that the decision trees first spectrally segment the images and then use the ancillary variables to separate spectrally similar classes. With such a structure, perhaps the decision trees have one branch for reference image data and other branches for cloud-filling data, and these branches span different signature ranges. Image normalization might make these branches less distinct by bringing their spectral ranges closer together, increasing spectral confusion. Though the normalizations reduce spectral differences between image dates, they may not reduce these differences enough to cause a complete merging of the different spectral hierarchies. In the end, the ancillary variables allow the decision trees to resolve the spectral confusion. As mentioned earlier, we know that the ancillary data reduce error from spectral confusion because they improve classification accuracy (from Kappa coefficients of 0.3–0.6 to values above 0.6 in this study). With spectral data only, some of the mosaic combinations lead to Kappa coefficients of agreement that are smaller for normalized imagery than for non-normalized or atmospherically corrected imagery. Because of the low accuracy level of all classifications of spectral data alone, however, this latter observation should be viewed with caution.

Ensuring that the classification training data represented the spatial and spectral range of each forest formation required some visual interpretation of the cloud-filled imagery. Imagery with cloud gaps filled from other image dates is more seamless for normalizations that use regression trees than for those that

Table 3. Per-class accuracies (user's and producer's) for selected decision tree classifications using one or two dates and no normalization method (none), linear regression normalization, and regression tree normalization.

	User's accuracy						Producer's accuracy					
	Dec. 1999			Dec. 1999 – Sept. 2002			Dec. 1999			Dec. 1999 – Sept. 2002		
Class	None	Linear regression	Regression tree	None	Linear regression	Regression tree	None	Linear regression	Regression tree	None	Linear regression	Regression tree
Bare	0.82	0.79	0.81	0.68	0.84	0.89	0.59	0.50	0.54	0.48	0.48	0.57
Urban	0.72	0.71	0.69	0.66	0.65	0.67	0.86	0.88	0.92	0.90	0.88	0.90
Sugar cane	0.63	0.66	0.68	0.66	0.67	0.62	0.71	0.81	0.78	0.62	0.77	0.70
Pasture	0.44	0.48	0.47	0.50	0.52	0.53	0.60	0.58	0.61	0.65	0.64	0.65
Drought deciduous woodland	0.56	0.54	0.59	0.51	0.56	0.53	0.42	0.48	0.46	0.54	0.67	0.60
Drought deciduous forest	0.87	0.87	0.90	0.90	0.89	0.89	0.63	0.63	0.65	0.65	0.61	0.63
Mixed coastal forest	0.85	0.82	0.85	0.85	0.89	0.90	0.71	0.72	0.79	0.69	0.69	0.81
Semideciduous forest	0.68	0.67	0.73	0.64	0.66	0.69	0.58	0.61	0.64	0.61	0.61	0.70
Evergreen forest	0.76	0.81	0.79	0.80	0.86	0.89	0.87	0.88	0.87	0.81	0.81	0.84
Cloud forest	0.93	0.92	0.92	0.92	0.92	0.92	0.89	0.89	0.88	0.89	0.88	0.86
Water	0.81	0.84	0.92	0.78	0.73	0.90	0.88	0.90	0.88	0.90	0.90	0.90

Note: The data suggest that slight increases in per-class accuracy are more common than decreases when classifications use two image dates and some type of normalization. However, these differences in per-class accuracies vary by class.

use linear regression, histogram matching, or image-based atmospheric correction. Consequently, normalizing image dates with regression trees can enhance visual interpretation. The procedure permitted us to develop fairly seamless cloud-filled imagery for each of two stages in vegetation phenology. One cloud-filled image has drought deciduous woody vegetation in a leaf-on stage, and the other has drought deciduous vegetation in a leaf-off stage. Together, these two cloud-filled images permitted us to collect a more comprehensive training dataset for classification.

The capacity to model complex nonlinear relationships probably explains why predicting a reference image from a subject image with regression tree models more closely matches imagery than do linear regression or histogram matching. The effects of phenological change on the relationships between bands from different dates can differ by class, which results in nonlinear relationships between those bands. The nonlinearity of regression trees, however, may make them inappropriate for use in the same way that others have normalized images in a time series for change detection. Regression tree models might tend to match changed areas to a reference date, which would hide land-cover change. In addition, the success of various normalization approaches probably depends on the extent of each class in the mutually cloud-free image parts that supply data for normalization models. Lastly, we observed during this and previous work (Helmer and Ruefenacht, 2005) that differences in sun–target–sensor geometry may negatively affect the success of regression tree normalization. When choices exist, image sets dated close to anniversary dates so that sun–target–sensor geometry is similar may yield better results.

Conclusions

When assembling cloud-filled imagery over Caribbean island landscapes, regression tree normalization of ETM+ bands 1–5 and 7, NDVI, and the band 4:5 ratio can produce more seamless imagery than image-based atmospheric correction or linear methods of radiometric normalization because it more closely matches vegetation phenology. Normalizing the data that fill cloud gaps may not affect classification accuracy when machine learning classifiers are used in combination with (*i*) training data that reflect class spectral and spatial variability (including in cloud-filled areas), and (*ii*) relevant ancillary data. At the same time, cloud-filled imagery assembled with no normalization can be difficult to visually interpret, which may hinder some applications. In this study, for example, fairly seamless cloud-filled imagery for more than one season was vital to distributing the visually interpreted training data over the extents of deciduous and semideciduous woody vegetation. In conclusion, when an application requires visual interpretation of satellite imagery, normalizing interdate image differences in cloud-filled imagery, particularly with regression trees, can be an important enhancement for Landsat ETM+ optical bands.

Acknowledgements

We thank the US AID-Caribbean Regional Programs Office, which funded the fieldwork for this research. Anonymous reviewers and Rachel Riemann made important manuscript comments. We also thank Justin Gray, Humfredo Marcano, Ariel Lugo, Tom Brandeis, Esther Rojas, Floyd Liburd, Lemuel Pemberton, Bernard Vanderpool, Ilis Palmer-Rannie, Rudy King, Ray Czaplewski, Michael Lefsky, Alex Hoppus, Mark Finco, Larry Tieszen, and Steve Schill for their contributions to completing this study. This research was conducted in cooperation with the University of Puerto Rico and the US Forest Service Rocky Mountain Research Station (RMRS).

References

Arroyo Mora, P., Sanchez-Azofeifa, G.A., Kalacska, M.E.R., Rivard, B., Calvo-Alvarado, J., and Janzen, D.H. 2005. Secondary forest detection in a Neotropical dry forest landscape using Landsat 7 ETM+ and IKONOS imagery. *Biotropica*, Vol. 37, No. 4, pp. 497–507.

Castro, K., Sanchez-Azofeifa, G., and Rivard, B. 2003. Monitoring secondary tropical forests using space-borne data: implications for Central America. *International Journal of Remote Sensing*, Vol. 20, No. 9, pp. 1853–1894.

Chavez, P.S. 1996. Image-based atmospheric corrections — revised and improved. *Photogrammetric Engineering & Remote Sensing*, Vol. 62, No. 9, pp. 1025–1036.

Cihlar, J., Ly, H., and Xiao, Q. 1996. Land cover classification with AVHRR multichannel composites in northern environments. *Remote Sensing of Environment*, Vol. 58, No. 1, pp. 36–51.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Management*, Vol. 20, pp. 37–46.

Condit, R., Watts, K., Bohlman, S.A., Pérez, R., Foster, R.B., and Hubbell, S.P. 2000. Quantifying the deciduousness of tropical forest canopies under varying climates. *Journal of Vegetation Science*, Vol. 11, No. 5, pp. 649–658.

Du, Y., Teillet, P.M., and Cihlar, J. 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment*, Vol. 82, No. 1, pp. 123–134.

Farr, T.G., and Kobrick, M. 2000. Shuttle Radar Topography Mission produces a wealth of data. *EOS Transactions of the American Geophysical Union*, Vol. 81, pp. 583–585.

FGDC. 1997. *National vegetation classification standard, Federal Geographic Data Committee, Vegetation Subcommittee*. Federal Geographic Data Committee (FGDC), US Geological Survey, Reston, Va. FGDC-STD-005.

Fiorella, M., and Ripple, W.J. 1993. Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogrammetric Engineering & Remote Sensing*, Vol. 59, No. 2, pp. 239–246.

Friedl, M.A., and Brodley, C.E. 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, Vol. 61, No. 3, pp. 399–409.

Helmer, E.H. 2000. The landscape ecology of secondary forest in montane Costa Rica. *Ecosystems*, Vol. 3, pp. 98–114.

Helmer, E.H. 2004. Forest conservation and land development in Puerto Rico. *Landscape Ecology*, Vol. 19, No. 1, pp. 29–40.

Helmer, E.H., and Ruefenacht, B. 2005. Cloud-free satellite image mosaics with regression trees and histogram matching. *Photogrammetric Engineering & Remote Sensing*, Vol. 71, No. 9, pp. 1079–1089.

Helmer, E.H., Cohen, W.B., and Brown, S. 2000. Mapping montane tropical forest successional stage and land use with multi-date Landsat imagery. *International Journal of Remote Sensing*, Vol. 21, No. 11, pp. 2163–2183.

Helmer, E.H., Ramos, O., Lopez, T.d.M., Quiñones, M., and Diaz, W. 2002. Mapping forest types and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot. *Caribbean Journal of Science*, Vol. 38, No. 3–4, pp. 165–183.

Holben, B.N. 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, Vol. 7, pp. 1417–1434.

Homer, C.G., Ramsey, R.D., Edwards, T.C., and Falconer, A. 1997. Landscape cover-type modeling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering & Remote Sensing*, Vol. 63, No. 1, pp. 59–67.

Homer, C.G., Huang, C., Yang, L., Wylie, B., and Coan, M. 2004. Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering & Remote Sensing*, Vol. 70, No. 7, pp. 829–840.

Howard, S.M., and Lacasse, J.M. 2004. An evaluation of gap-filled Landsat SLC-off imagery for wildland fire burn severity mapping. *Photogrammetric Engineering & Remote Sensing*, Vol. 70, No. 4, pp. 877–880.

Huang, C., and Townshend, R.G. 2003. A stepwise regression tree for nonlinear approximations: applications to estimating subpixel land cover. *International Journal of Remote Sensing*, Vol. 24, No. 1, pp. 75–90.

Kappelle, M., Kennis, P.A.F., and de Vries, R.A.J. 1995. Changes in diversity along a successional gradient in a Costa Rican upper montane *Quercus* forest. *Biodiversity and Conservation*, Vol. 4, pp. 10–34.

Leica Geosystems GIS & Mapping, LLC. 2003. *ERDAS field guide*. Leica Geosystems GIS & Mapping, LLC, Atlanta, Ga.

Lo, T.H.C., Scarpace, F.L., and Lillesand, T.M. 1986. Use of multitemporal spectral profiles in agricultural land-cover classification. *Photogrammetric Engineering & Remote Sensing*, Vol. 54, No. 4, pp. 535–544.

Lugo, A.E., Schmidt, R., and Brown, S. 1981. Tropical forests in the Caribbean. *Ambio*, Vol. 10, pp. 318–324.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Ke, J. 2000. Biodiversity hotspots for conservation priorities. *Nature (London)*, Vol. 403, pp. 853–858.

Olsson, H. 1993. Regression functions for multitemporal relative calibration of Thematic Mapper data over boreal forest. *Remote Sensing of Environment*, Vol. 46, No. 1, pp. 89–102.

Pax-Lenney, M., Woodcock, C.E., Macomber, S.A., Gopal, S., and Song, C. 2001. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment*, Vol. 77, No. 3, pp. 241–250.

Richards, J.A. 1993. *Remote sensing digital image analysis, an introduction*. Springer-Verlag, New York.

Rodrigues, A.S.L., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Cowling, R.M., Fishpool, L.D.C., da Fonseca, G.A.B., Gaston, K.J., Hoffmann, M., Long, J.S., Marquet, P.A., Pilgrim, J.D., Pressey, R.L., Schipper, J., Sechrest, W., Stuart, S.N., Underhill, L.G., Waller, R.W., Watts, M.E.J., and Yan, X. 2004. Effectiveness of the global protected area

network in representing species diversity. *Nature (London)*, Vol. 428, No. 6983, pp. 640–643.

Scaramuzza, P., Micijevic, E., and Chander, G. 2004. *SLC gap-filled products, Phase 1 methodology.* Available from http://landsat.usgs.gov/data_products/slc_off_data_products/documents/SLC_Gap_Fill_Methodology.pdf [cited 6 August 2007].

Schott, J.R. 1997. *Remote sensing: the image chain approach*. Oxford University Press, New York.

Schott, J.R., Salvaggio, C., and Volchok, W.J. 1988. Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment*, Vol. 26, No. 1, pp. 1–14.

Schriever, J.R., and Congalton, R.G. 1995. Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the Northeast. *Photogrammetric Engineering & Remote Sensing*, Vol. 61, No. 3, pp. 321–327.

Scott, J.M., Davis, F., Csuti, G., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards, T.C., Jr., Ulliman, J., and Wright, R.G. 1993. Gap analysis: a geographical approach to protection of biological diversity. *Wildlife Monographs*, Vol. 123, pp. 1–41.

Seto, K., and Liu, W. 2003. Comparing ARTMAP Neural Network with the maximum-likelihood classifier for detecting urban change. *Photogrammetric Engineering & Remote Sensing*, Vol. 69, No. 9, pp. 981–990.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., and Macomber, S.A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*, Vol. 75, No. 2, pp. 230–244.

Vogelmann, J.E. 1988. Detection of forest change in the Green Mountains of Vermont using multispectral scanner data. *International Journal of Remote Sensing*, Vol. 9, No. 7, pp. 1187–1200.

Vogelmann, J.E., Howard, S., Yang, L., Larson, C., Wylie, B., and Van Driel, N. 2001. Completion of the 1990s National Landcover Dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering & Remote Sensing*, Vol. 67, No. 6, pp. 650–661.

Wolter, P.T., Mlandenoff, D.J., Host, G.E., and Crow, T.R. 1995. Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery. *Photogrammetric Engineering & Remote Sensing*, Vol. 61, No. 9, pp. 1129–1143.

Woodcock, C.E., Macomber, S.A., Pax-Lenney, M., and Cohen, W.B. 2001. Monitoring large areas for forest change using Landsat: generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, Vol. 78, No. 1–2, pp. 194–203.