



This is a peer-reviewed, final published version of the following document and is licensed under Creative Commons: Public Domain Dedication license:

**Goodenough, Anne E. and Goodenough, Andrew S. (2011) Development of a Rapid and Precise Method of Digital Image Analysis to Quantify Canopy Density and Structural Complexity. ISRN Ecology, 2012. pp. 1-11. ISSN 2090-4622**

Official URL: <http://dx.doi.org/10.5402/2012/619842>

DOI: <http://dx.doi.org/10.5402/2012/619842>

EPrint URI: <http://eprints.glos.ac.uk/id/eprint/3352>

#### **Disclaimer**

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

## Research Article

# Development of a Rapid and Precise Method of Digital Image Analysis to Quantify Canopy Density and Structural Complexity

Anne E. Goodenough<sup>1</sup> and Andrew S. Goodenough<sup>2</sup>

<sup>1</sup>Department of Natural and Social Sciences, Francis Close Hall, University of Gloucestershire, Swindon Road, Cheltenham, Glos GL50 4AZ, UK

<sup>2</sup>St Briavels, Forest of Dean, Gloucestershire, UK

Correspondence should be addressed to Anne E. Goodenough, aegoodenough@glos.ac.uk

Received 21 October 2011; Accepted 4 December 2011

Academic Editors: A. Chappelka and M. Rossetto

Copyright © 2012 A. E. Goodenough and A. S. Goodenough. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Estimation of canopy density is necessary for ecological research and woodland management. However, traditional manual methods are time consuming and subject to interobserver variability, while existing photographic methods usually require expensive fish-eye lenses and complex analysis. Here we introduce and test a new method of digital image analysis, CanopyDigi. This allows user-defined threshold to polarise the 256 grey shades of a standard monochrome bitmap into dark “canopy” and light “sky” pixels (the threshold being selected using false-colour images to ensure its suitability). Canopy density data are calculated automatically and rapidly, and, unlike many other common methods, aggregation data are obtainable using Morisita’s index to differentiate closed (diffuse light) and open (direct light) canopies. Results were highly repeatable in both homogeneous and heterogeneous woodland. Estimates correlated strongly with existing (nondigital) canopy techniques, but quicker and with significantly lower interobserver variability (CV = 3.74% versus 20.73%). We conclude that our new method is an inexpensive and precise technique for quantifying canopy density and aggregation.

## 1. Introduction

Vegetation structural complexity comprises several parameters including the density and height of different vegetation layers, the percentage coverage of each layer, and species composition [1]. In a woodland ecosystem, one of the most important aspects of overall structural complexity is the canopy density and aggregation [2, 3]. These parameters influence stand productivity and species composition, as well as the density and growth of the understorey, which is a key determinant of habitat [4]. Estimation of canopy coverage is important in studies of species-habitat relationships (e.g., [5–10]) and also for land managers given the increasing emphasis on sustainable woodland management. The effective use of canopy data in ecological research, or to inform woodland management, demands that measurements are accurate, precise, and obtainable using a quick, easy, and inexpensive method [11]. Estimates also need to have a high consistency and low interobserver variability to avoid generation of erroneous results [12–14].

*1.1. Canopy Measurements: Coverage and Density.* Somewhat surprisingly given its importance, there is no standard method of measuring canopy on a relatively small scale (e.g., for specific woodland plots). Jennings et al. [15] distinguish two basic types of measurement of forest canopies; the percentage of canopy coverage—the area of ground covered by a vertical projection of the canopy—or the percentage of canopy density (also referred to as closure)—the area of ground covered by canopy when viewed from a single point. Methods to assess these will be considered briefly below (for a more in-depth review, see [15, 16]).

*1.1.1. Coverage.* The most straightforward method of estimating coverage is by eye. This can be estimated as a percentage or using a scale such as the generic DOMIN or Braun-Blanquet [17, 18]. These are quick and simple, but are also subjective and very variable [19]. Indeed more variability in visual estimates can be attributed to the evaluators than to actual differences in the vegetation cover [19]. To increase objectivity and decrease variability, canopy

cover is often assessed using a sighting tube [16]. These either have a single central cross-wire or an acetate grid with several intersections, and recorders look up through the tube and note if the cross-points are aligned with vegetation or sky. The alignment ratio becomes the coverage estimate [6, 13]. Unfortunately numerous measurements are recommended at each sample location, and it is very difficult to keep the sighting tube vertical, which is essential to avoid overestimating the canopy [20].

*1.1.2. Density.* Canopy density is more closely linked to microclimate and light regime than canopy cover and is therefore generally preferred in ecological studies [16, 21]. It can be measured visually using the specific crown illumination index [22], using a spherical densiometer [23] or using a moosehorn [24]. A spherical densiometer is a mirrored dot-etched hemispherical dome. When placed under the canopy, the proportion of dots covered by the canopy's reflection estimates its density. As well as this being expensive and specialised [4], there are concerns about its accuracy [15, 25] and it is subject to high interobserver variability and, therefore, a lack of precision [6, 26]. The moosehorn is a transparent dot-marked screen that held slightly away from the eye; the ratio of vegetation and sky alignments are recorded as before. Not only is this method cumbersome and the equipment fragile, but it also underestimates canopy density as approximately 23% of canopy gaps remain outside the field of view [12, 14].

It should also be noted that there is a considerable difference between a canopy with many small gaps, which allow diffuse light to penetrate (Figures 1(a) and 1(b)), and a canopy with one or more large gaps, which allow direct light to penetrate (Figures 1(c) and 1(d)) [27]. However, none of the above methods can discern this objectively, such that this important habitat parameter is disregarded in the vast majority of ecological studies. Indices to quantify structural complexity to distinguish different types of vegetation aggregation do exist, for example, gap fraction estimates, leaf area index, canopy openness, and sky factor [28, 29]. However, these are only obtainable using complex methods such as: (1) ground-level tripod-mounted scanning using light detection and ranging (lidar) units or terrestrial laser scanners, which can digitise structural information about the canopy very rapidly [28, 30]; or (2) canopy photography, both of which are often prohibitively expensive.

*1.2. Canopy Photography.* Determination of canopy density from photographs via estimation of light penetration has been suggested previously for canopy estimation [4, 31, 32]. Usually, this refers to hemispherical photography [33, 34], which gives an almost complete view of the canopy, and analysed using specially developed software such as Gap Light Analyzer [27] and HemiView [35]. However, there are several drawbacks to the use of hemispherical photography to quantify canopy parameters. The most important issue is the high cost (and thus low accessibility) of equipment [16], but hemiplots (images resulting from hemispherical photography) also: (1) include all vegetation present, including understorey and field layers, which may not be desirable, (2) cannot be taken at ground level on slopes because the slope itself is

photographed, and (3) require complex spherical geometric calculations as there is no direct relationship between canopy area and its photographic representation [36].

As a result of these limitations, particularly that of cost, canopy photography is often dismissed as being complex, time consuming, and expensive [12, 15] and there have been calls for the development of other methods that are cheaper and easier to use [16]. If tree canopy is the only structural complexity variable of interest, standard digital photographs (rather than hemiplots) can be taken using a standard camera and analysed much more easily. Some work on computer analysis of standard canopy photographs has been undertaken previously. Engelbrecht and Herz [21] examined canopy photographs using Adobe Photoshop (San Jose, USA) whereby a pure black-and-white image was created by maximising the image's contrast, and the percentage of black (canopy) pixels was determined. This is a straightforward technique that uses commercial software. However, photographs need to be taken under uniformly overcast skies, which places considerable (often insurmountable) constraints on fieldwork. Moreover, although the method worked well on equatorial evergreen canopies, by the authors' own admission, it does not appear to be transferable to other situations. Indeed, even when standard photographs of an artificial tree taken under laboratory conditions (uniform background, optimal lighting, etc.) were processed in this way, the resultant images were not a particularly accurate representation of the true conditions, with small areas of tree being regarded as background [28]. Purcell [37] and Richardson et al. [38] have used a conceptually similar approach to quantify ground vegetation cover from colour photographs by counting the number of green pixels within specific hue and saturation ranges. Again, however, there were constraints: plants in shadow were often not included, sun flecks could bias analysis, and the technique only worked with green vegetation (such that it would not work, for example, for trees in the autumn).

In this study, we introduce a new photographic method to assess percentage density via estimation of light penetration. This involves flexible digital image analysis of standard canopy photographs taken with a cheap digital camera (not hemiplots) with a user-selected threshold between dark (canopy) and light (sky) areas. Image analysis also includes quantification of canopy aggregation (patchiness) in order to quantify this important variable objectively. We test the new technique in two different situations (homogeneous and heterogeneous woodland) and then compare it to existing methods of quantifying canopy coverage and density in order to evaluate relative levels of interobserver variation. Finally, we consider the wider applicability of the new technique for surveying other aspects of vegetation structural complexity.

## 2. Materials and Methods

The following details conversion of monochrome canopy photographs to false-colour images using a series of threshold values to differentiate between dark "canopy" and light "sky" pixels to thus obtain information on canopy density (percentage quantity derived from estimation of light

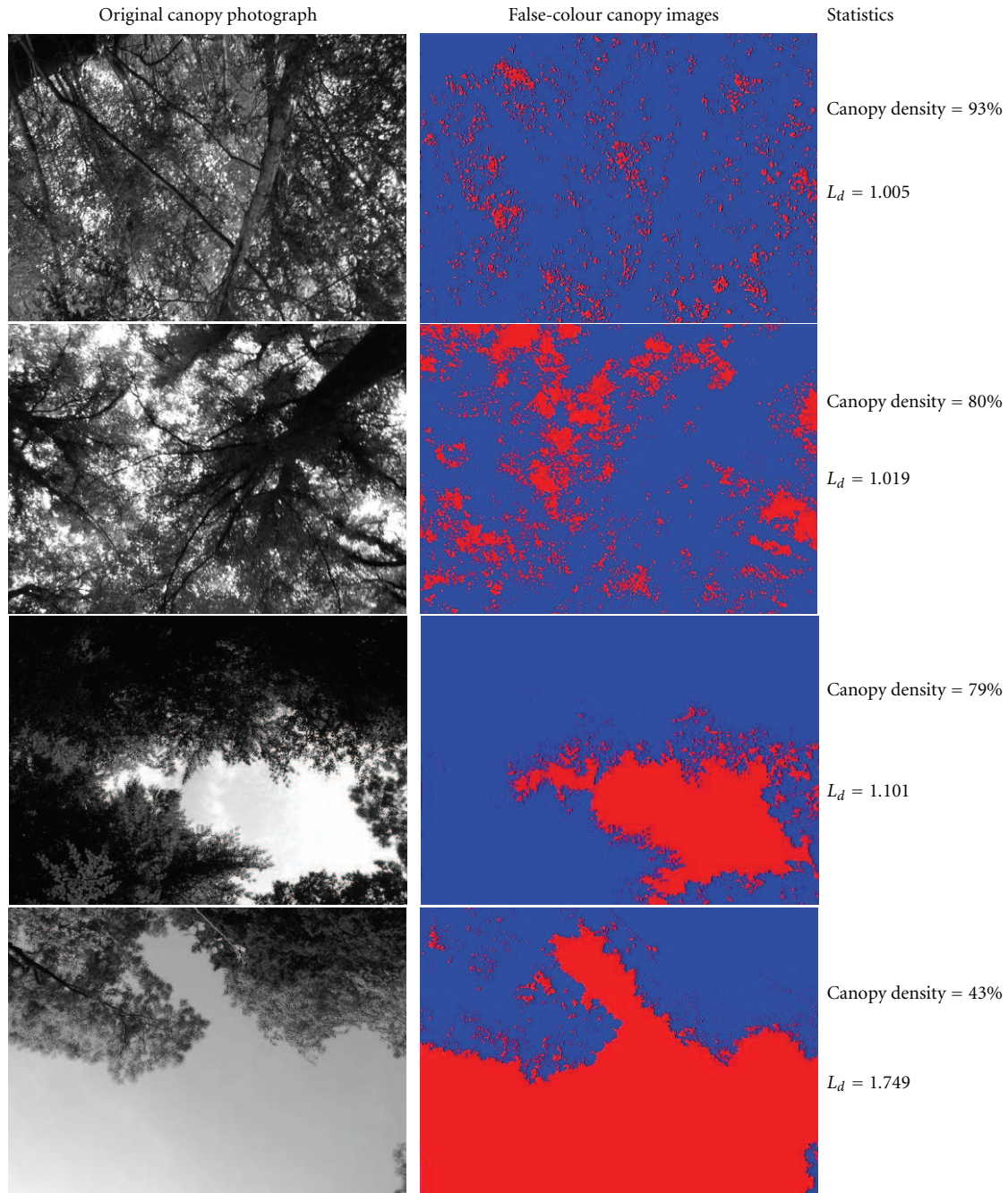


FIGURE 1: Canopy photographs and their respective false-colour images where dark “canopy” pixels are coloured blue (light grey in printed paper) and light “sky” pixels are coloured red (dark grey in printed paper). Pixels are assigned on the basis of a threshold between greyscale values 1 and 255 (all the above have been created using threshold value 155; see Section 2). Density data (percentage) and aggregation (Morisita’s index) are given.

penetration) and aggregation (patchiness). Methods used to test the technique are then outlined. All aspects of determination and quantification are undertaken using a suite of computer programs written in Microsoft Visual Basic. This suite of programs, CanopyDigi, is included with this paper as electronic supplementary material, together with a user-friendly interface, full documentation, and help files.

*2.1. Developing the Digital Image Analysis Technique.* A threshold approach was used to obtain estimates of canopy density by digital analysis of canopy photographs held as monochrome bitmap (BMP) images. This made use of the fact that within any monochrome BMP, there is provision for 256 shades of grey, ranging from black (0) to white (255). Although in principle a canopy estimate could be obtained simply by counting the number of pixels in each

grey shade using a demarcation value (threshold) at and above which pixels are regarded as canopy (an approach originally suggested by Anderson) [39], this was found to be problematic. Using a fixed threshold (e.g., [40]) placed too many restrictions on fieldwork to be practicable, while manual selection of a threshold was entirely subjective when there was no way of checking the demarcation value objectively [4, 11]. As preliminary testing showed that the most appropriate threshold varied significantly according to tree type, ambient light, and time of day, it was decided to undertake analyses on specially created false colour images with all “canopy” pixels being coloured blue and all “sky” pixels being coloured red. In this way, the suitability of any given threshold could be examined in a “determination” stage and canopy then analysed using that threshold in a “quantification” stage.

Theoretically, it should be possible to automate the determination stage, either fully or partially, using threshold-selection algorithms, and there have been several attempts to do this (e.g., [41, 42]). However, these algorithms are not always successful in determining an appropriate threshold [43], especially when the relative contrast between vegetation and sky changes across the image (e.g., sun-illuminated vegetation against white cloud in one part of the image and shaded vegetation against bright blue sky in another part of the image) [44] and add substantially to processing time. A quick examination of photograph-specific histograms demonstrates the problem. If canopy photographs were always bimodal in greyscale pixel frequency (i.e., if canopy equalled black and sky equalled white), it would be possible to choose either: (1) the midpoint between the two modes; or (2) the midrange value with the lowest pixel occurrence and make this the automated threshold point. Alternatively, if the histograms were always normally distributed, it would be possible to ensure that all values below the mode were treated as canopy and all values above this were treated as sky. Although both of these approaches would be easy to code, extensive testing showed that these approaches were, in reality, ineffective. Histograms were very location specific, some being near normal, others being bi- or multimodal, and others tending towards a uniform, U-shaped, or Poisson distribution (Figure 2). We therefore adopted a user-controlled determination stage as detailed above. This approach also has the huge advantage that the user can see the “fit” of the false-colour image to the original image. This means that if it is simply not possible to have one threshold to separate sky and canopy across the entire image (i.e., if a variable, context-specific, threshold is most appropriate, an approach which has been showcased for individual photographs but has yet to be developed into batch-processing software [44]) it is possible for the user to interpret canopy statistics for that image with caution. In other words, uncertainly can be recognised and allowed for explicitly rather than being hidden in the realms of computer code. Usefully, however, based on the results of testing the technique as reported here, we found that one threshold from a restricted (standardised) subset of options was appropriate in the majority of situations and that this could be used in the first (and often only) run through the determination stage (see Results). This has the twin advantages of making the

threshold, if not automated, at least easy to select and quick to implement in the majority of situations, while retaining complete flexibility to enable analysis of atypical images.

*2.1.1. Determination Stage.* To determine the most appropriate threshold for any picture, eight false-colour images, each at a different default threshold, were created. The most appropriate threshold (i.e., that which created the “best fit” red/blue image) was then selected following simultaneous visual comparison of each false-colour image with the original photograph. In order to minimise the amount of time during which user input was required, batch processing was used whereby multiple photographs were run through the determination stage, without the need for user input. The most appropriate false colour image for each photograph was user selected in a rapid post hoc selection process. It should be noted that although the eight standard (default) thresholds were determined following extensive testing, the ability for users to change the preset thresholds for any pixel shade between 1 and 254 was retained. This ensured that the digital image analysis was entirely flexible [21], such that atypical photographs (e.g., photographs with deep shadows or light flecks) were still analysable. Importantly, the threshold value itself was immaterial as long as the false-colour image distinguished correctly between canopy and noncanopy areas of the photograph, such that photographs processed using different thresholds were still directly comparable.

*2.1.2. Quantification Stage.* To provide an estimate of canopy density for each false-colour image selected during the determination stage, the percentage of blue (canopy) pixels was calculated. Then, to provide an estimate of canopy dispersion, Morisita’s index, one of the most robust distribution measures [45], was quantified. To do this, each image was electronically divided into 12 subimages (equivalent to the plots or quadrates used in field applications of Morisita’s index (e.g., [46])), and the number of canopy pixels was calculated for each subpicture. The index was then calculated using

$$I_d = n \left( \frac{\sum X^2 - N}{N(N-1)} \right), \quad (1)$$

where  $I_d$  is Morisita’s index of dispersal,  $n$  is the total number of subpictures,  $N$  = total number of canopy pixels in the entire image and  $X^2$  is the square of the canopy pixels in each subpicture ( $X_1^2, X_2^2$ , etc.). The index value increased with increasingly large gaps in the canopy, producing an interval variable for subsequent analysis.

Batch processing without the need for user input was used, and the final density and Morisita statistics for all original photographs in one batch were written out to a single CSV (comma-separated variable) file to facilitate further analysis (in Excel, SPSS, Minitab, etc.), as well as to individual photograph-specific text files.

*2.2. Testing the Digital Image Analysis Technique.* To assess the effectiveness of the digital image analysis technique, canopy photographs were taken at Nagshead Nature Reserve

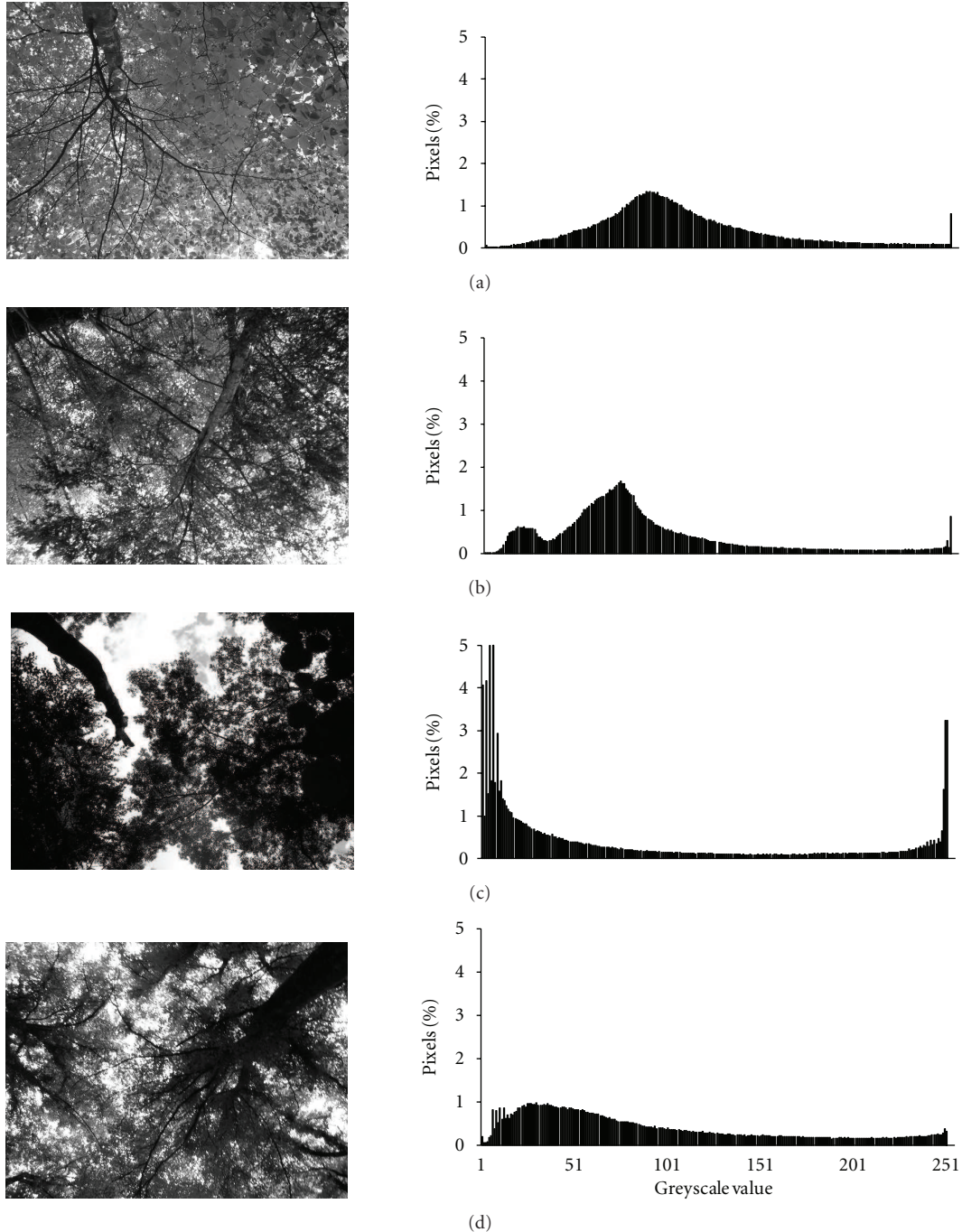


FIGURE 2: Canopy photographs and their greyscale pixel histograms showing differing distributions.

(Gloucestershire, UK) and Lady Park Wood (Monmouthshire, UK) ( $n = 295$  and  $58$ , resp.). Nagshead is a homogeneous wood, dominated by mature Pedunculate oaks (*Quercus robur*), which are around 200 years old and have an average height of about 25 m. Lady Park Wood is a heterogeneous, nonintervention, wood with a mixed-aged, mixed-height, canopy (60–200 years old; 15 m–30 m + in height) comprising oak (*Q. robur*), ash (*Fraxinus excelsior*), elm (*Ulmus glabra*), beech (*Fagus sylvatica*),

small-leaved lime (*Tilia cordata*), and yew (*Taxus baccata*). To obtain images suitable for analysis (i.e., that captured the necessary information but were small enough to ensure rapid computer analysis), photographs were taken vertically upwards using a digital camera (Minolta Dimage Z1) set to monochrome format with a standardised image size of  $640 \times 480$  pixels (following [37]) (note that photographs could equally well have been converted to these settings after fieldwork using standard software such as Paintshop

Pro or Matlab, or using shareware such as ReaConverter (<http://www.reaconverter.com/>), which all use batch processing to minimise the need for human interaction, and all of which would also convert photographs of a different aspect ratio than 1 : 1.333, if necessary). The resolution of 640 \* 480 was used because any higher resolution would not improve the accuracy of the technique, merely generating spurious precision and substantially increasing computer processing time; larger files can easily be converted (see above). The focal length at which a canopy photograph is taken will influence results since the focal length affects the actual area of the canopy being photographed (taking one image at, for example, 18 mm, and another at 50 mm would result in very different images, and, thus, in all likelihood, different canopy statistics would be obtained after analysis). There is no “right” focal length but if images need to be directly comparable—for example, in longitudinal monitoring at the same locations—it would be necessary to be standardised on one agreed focal length. In this case, focal length was set to 35 mm, the focus was set to infinity, and aperture and shutter speed were determined automatically. To avoid perspective influencing the canopy calculations, the verticality of the photographs was assured using a two-way spirit level affixed to the back of the camera. Following fieldwork, the images were converted from lossy JPEG format (the standard digital camera format) to lossless BMP format using ReaConverter (see above) to ensure that all photographs had an identical number of pixels, thereby facilitating analysis.

*2.3. Threshold Sensitivity Analysis.* To determine the effect of selecting a nonideal threshold on the resultant canopy density percentage and Morisita’s index data, a subset of 20 photographs, all taken in Lady Park Wood by one individual (AEG) and all judged to have an ideal threshold of 155, were selected randomly from the 54 such photographs matching these criteria. These were run through CanopyDigi at a threshold of 155, and four thresholds either side of this in increments of five greyscale values (i.e., 135, 140, 145, 150 and 160, 165, 170, and 175). The canopy percentage estimates were extracted for each threshold on a per-photograph basis and were run through a repeated measures ANOVA (after allowing for sphericity using the Greenhouse-Geisser correction) whereby each photograph (site) was treated as an individual case on the basis that nonsignificance would be good evidence that the technique is robust to relatively substantial deviations from the most suitable threshold. The same process was used for the Morisita values.

*2.4. Comparison of the Digital Analysis Technique Using Photographs from Different Cameras.* Although there is no reason to suppose that two images, taken consecutively at the same place and the same focal length but with different cameras, would give differing results, we tested this empirically. One of us (A. E. Goodenough) took one canopy photograph at each of 12 different locations with the camera used previously (a compact digital set to 35 mm) and, immediately thereafter, took a second photograph with a single lens reflex (SLR) camera (Canon 450D mounted with an 18–55 mm lens set to 35 mm equivalent). Photographs were taken under

field conditions and without the aid of a tripod. We ran all 24 images through CanopyDigi and tested for differences in the recorded canopy density (percentage) and dispersion (Morisita’s index) using separate paired-samples *t*-tests (one for each of the two canopy parameters; percentage data were arcsine square root transformed to normalise them).

*2.5. Comparing Methods: Interobserver Variation and Sampling Effort.* To compare interobserver variation of canopy photographs with existing methods (1 = visual estimation, 2 = cross-wire sighting tube, and 3 = moosehorn), ten volunteers were asked to record the canopy at ten sample locations. All volunteers, drawn from both sexes and a variety of age ranges, were novices—they had no experience of taking canopy photographs or analysing the resultant images, nor any experience of the traditional methods used to estimate canopy. In each case a 5 m<sup>2</sup> plot was laid to one side of a central tree. To prevent the use of other techniques influencing visual estimates, these were made first [47], with the remaining methods being employed in a random order [6]. Visual estimation (a percentage estimate to the nearest 5%), the moosehorn estimation, and the digital photograph were taken from the centre of the plot. To obtain the recommended 100 cross-wire sighting tube readings [20], the plot was conceptually divided into 100 squares (10 × 10), and a measurement was taken in the middle of each small square, the ratio of vegetation to sky alignments providing the coverage estimation. To eliminate potentially confounding variables, all fieldwork was undertaken in comparable lighting and weather conditions and within a 48-hour period. At each sample location, the volunteer took two photographs of the canopy. The first photograph was always analysed unless there was a problem (e.g., camera shake) when the second photograph was used instead [32]. The camera and settings were consistent and as described above. To ensure that the volunteers were not influenced by the estimates given by others, each completed their fieldwork in isolation [6]. The coefficient of variation (CV) was calculated for each sample location using the estimate of canopy density from each observer ( $n = 10$ ) and expressed as a percentage ( $CV = (\text{standard deviation}/\text{mean}) * 100$ ). Significant differences in the variability of the estimates derived using digital image analysis versus the other techniques were ascertained using a series of three *F* tests. In all cases, square of the differences in all 10 canopy estimates (one at each sample location) was calculated on a per-individual basis to remove location-specific effects (as per [38]). The mean canopy calculation for each sampling site, according to each method, was calculated so that methods could be compared using correlation analysis. The sampling effort, a variable that is often ignored when canopy quantification methods are compared, was also recorded by quantifying the mean time taken to achieve estimates by each method, both in terms of fieldwork and postfieldwork calculations.

### 3. Results

*3.1. Testing the Digital Image Analysis Technique.* No problems were experienced in taking the canopy photographs.

TABLE 1: Interobserver variability in quantifying canopy density at 10 different sample locations using three traditional methods and digital image analysis.

Method	Difference between highest and lowest canopy estimates at each sample location (%points)			Coefficient of Variation (%)		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
Visual estimation	28	75	50	11.63	50.92	26.37
Cross-wire tube	29	62	42	16.58	27.38	18.86
Moosehorn	30	62	39	11.13	26.97	16.96
Digital image analysis	6	17	10	2.97	6.66	3.74

As regards processing, the same threshold value (greyscale value 155) was suitable for the majority of photographs (all 295 photographs taken at Nagshead Nature Reserve, the homogeneous site, and 54 of the 58 photographs taken at Lady Park Wood, the heterogeneous site). Of the Lady Park Wood photographs, two required a lower threshold (greyscale value 125) due to dark blue sky being regarded initially as canopy, while the other two needed a higher threshold (greyscale value 185) due to sun flecks on the vegetation being regarded initially as sky (this latter could potentially be minimised in camera with the use of a polarising filter attached for a single photograph that could then be processed as normal, or, alternatively, by taking one photograph with a polariser and one without a polariser and combining the images electronically to add in the sunlight-reflected pixels before use of CanopyDigi). The percentage data gave information on canopy density itself, while Morisita's index quantified the aggregation of canopy gaps in each photograph (Figure 1). Digital analysis of the test photographs resulted in a wide range of percentage density estimates (minimum = 26%; maximum = 93%). Morisita's index values provided an interval variable for subsequent analysis and ranged from 1.002 to 2.745.

**3.2. Threshold Sensitivity Analysis.** When 20 photographs, all with an ideal threshold of 155, were run through CanopyDigi at lower and higher thresholds (135, 140, 145, 150; 160, 165, 175, 175, resp.), neither the canopy percentages nor the Morisita values calculated at these different thresholds differed significantly from one another on a per-photograph basis (Greenhouse-Geisser repeated measures ANOVA:  $F = 6.279$ , d.f. = 1,  $P = 0.085$  and  $F = 1.266$ , d.f. = 1.022,  $P = 0.324$ , resp.). This suggests that, while it is obviously best to select the ideal threshold to best distinguish canopy and sky areas of the photograph, selecting a threshold quite substantially lower or higher does not significantly confound the results.

**3.3. Comparison of the Digital Analysis Technique Using Photographs from Different Cameras.** There was no difference in either canopy density (percentage) or dispersion (Morisita's index) estimates for the 12 locations photographed pairwise using two different cameras; the standard compact camera described above and as a digital SLR camera (paired-samples  $t$ -tests  $t = 0.95$ , d.f. = 11,  $P = 0.341$  and  $t = 0.389$ , d.f. = 11,  $P = 0.705$ , resp.).

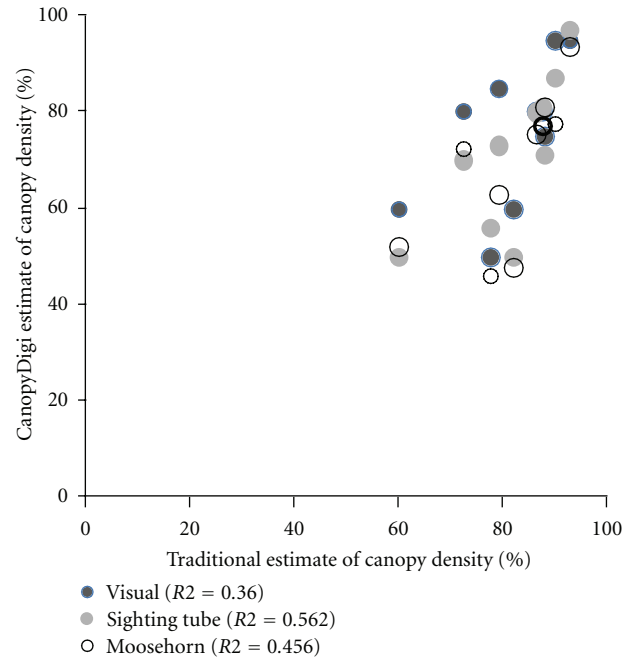


FIGURE 3: Mean estimates of canopy density from digital image analysis correlated with estimates made at the same locations visually, with a sighting tube and using a moosehorn.

**3.4. Comparing Methods: Interobserver Variation and Sampling Effort.** Mean canopy density estimates derived using digital analysis correlated strongly and significantly with mean estimates from each of the traditional methods (correlation of photographically derived estimates with visual estimates  $r = 0.600$ ,  $n = 10$ ,  $P = 0.033$ ; with sighting tube estimates  $r = 0.750$ ,  $n = 10$ ,  $P = 0.006$ ; with moosehorn estimates  $r = 0.680$ ,  $n = 10$ ,  $P = 0.015$ ; Figure 3), indicating that the methods themselves were comparable. However, digital estimates had much lower interobserver variability than estimates made using traditional methods (Table 1; Figure 4). The mean difference between the highest and lowest canopy estimates at each sample location was 10 percentage points for digital image analysis and 44 percentage points for all other methods combined, with CV values being 3.74% and 20.73%, respectively. When compared statistically, canopy calculation using digital image analysis had significantly lower interobserver variability than visual estimates ( $F = 147$ ; d.f. = 9;  $P < 0.001$ ), estimates



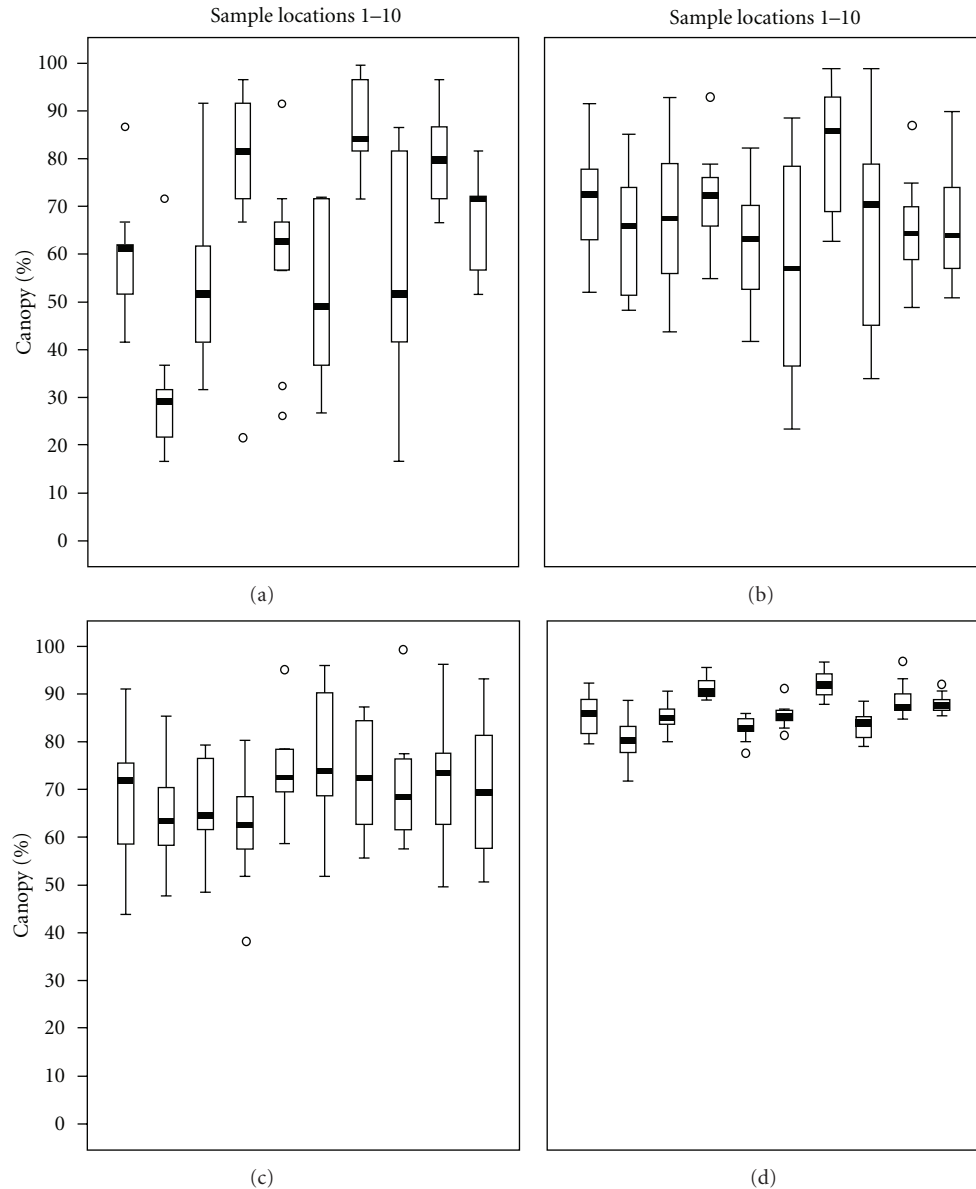


FIGURE 4: Interobserver variability in canopy estimates of ten observers surveying ten sample locations: (a) visually, (b) with a sighting tube (100 observations per person per sample location), (c) using a moosehorn, and (d) using digital image analysis. The long vertical line shows the 95% trimmed range, the box shows the interquartile range, the solid black horizontal lines show the median, and open circles show outliers ( $>1.5$  standard deviations from the mean).

derived from sighting tubes ( $F = 81$ ; d.f. = 9;  $P < 0.001$ ), and estimates derived from using the moosehorn ( $F = 65$ ; d.f. = 9;  $P < 0.001$ ). Indeed, the digital analysis method produced results with significantly lower variability even for the sample location (site 5) that had the smallest difference in interobserver variability between estimates made using the digital technique compared with those made using visual, sighting tube and moosehorn methods ( $F = 56, 41$  and  $17$ , respectively; d.f. = 9 and  $P < 0.001$  in all cases).

As regards sampling effort, visual estimation was the quickest method, requiring just 20 seconds per sampling location to record the information and later computerise it (Figure 5). Conversely, the sighting tube methods were

very time consuming. The only way to achieve an acceptable sampling speed was to use two people, one taking the measurements, the other recording them. Even then, to survey each site took an average of 6.5 minutes in the field and a further 1.5 minutes to calculate the ratio of vegetation to sky alignments. The moosehorn method was quick in the field—about 30 seconds—but quite time consuming after fieldwork, taking an average of 3 minutes per site to process information. The photographic method was extremely quick in the field: just 10 seconds including camera set-up time. Because of automation and batch processing restricting user input in the determination stage to selecting the most appropriate threshold value for each picture, the actual time

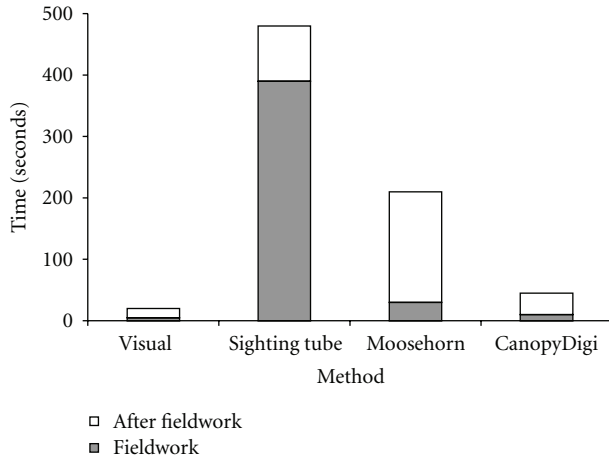


FIGURE 5: The sampling effort per sample location of canopy estimation using three traditional methods and digital image analysis, divided into time spent undertaking fieldwork after fieldwork analysis. Timings are means based on 10 sample locations sampled by 10 observers each (total per method  $n = 100$ ). Note that >85% of the after fieldwork processing of digital image analysis was computer processing during which no user interaction was necessary, while for all other methods total time and user interaction time are synonymous. The computer processing time for digital image analysis would increase a little if files need to be converted into 640\*480 greyscale BMP format from other file types, but this would not increase the amount of time for which interaction was necessary (due to automated batch processing).

spent at the computer to achieve objective quantification of canopy density and dispersion at all 353 sites was about 30 minutes, or just over 5 seconds per site (the time needed to choose the correct threshold for each photograph). Computer processing time, during which no user input was necessary (to generate eight individual false-colour pictures for each photograph and to calculate statistical information from each selected false-colour picture), was about 3 hours, or around 30 seconds per picture—these processes were run overnight. Thus the overall time taken to complete the photographic after fieldwork analysis was 35 seconds per picture, the vast majority (>85%) of which was computer processing time during which no human intervention was necessary.

#### 4. Discussion

CanopyDigi is a precise method of determining canopy density and enables objective mathematical calculation of canopy aggregation, providing useful quantification of canopy structure and differentiating between canopies with many small gaps, which allow diffuse light to penetrate (Figures 1(a) and 1(b)), and canopies with one or more large gaps, which allow direct light to penetrate (Figures 1(c) and 1(d)). The digital analysis method is much cheaper than advanced lidar scanning [28, 30] or hemispherical photography approaches [33, 34] and is comparable with existing commonly used techniques such as the use of a sighting tube or moosehorn, as demonstrated by a series of strong correlation coefficients. It should also be noted that previous work

has suggested that analysis of standard canopy photographs correlates strongly with technology-heavy and expensive lidar scanning [28]. If the camera used an integrated global positioning system (GPS) functionality, or a GPS attachment is used, it would also be possible to create interactive maps of the canopy images tagged with the statistics produced by CanopyDigi (in a similar way to Stafford et al. [48]).

*4.1. Comparing Methods: Interobserver Variation and Sampling Effort.* Low levels of interobserver variability are important in all studies using data collected by multiple individuals. Use of digital image analysis provides canopy estimates that have a much lower level of interobserver variation than what occurs using traditional (visual, sighting tube, and moosehorn) methods, both overall and for every individual sample location. Nine of the ten observers found estimating coverage visually difficult, with most finding it especially hard for sample locations dominated by coniferous species (locations 6 and 8). Use of the sighting tube was found to be difficult on sloping ground, and constant use was found to be fatiguing. When using the moosehorn, any small movement of the canopy made it very challenging to assess whether vegetation was aligned with the dots, decreasing the consistency of the techniques between observers. Conversely, most people found the camera method comparatively easy, and the precaution of each volunteer taking two photographs at each sample location proved almost unnecessary as the first photograph was suitable for analysis (no camera shake) in all but one case out of 100.

In terms of sampling effort, assessing canopy visually is rapid and gives an immediate percentage estimate with no post-fieldwork effort; however, the precision and consistency of the estimates is very poor. The sighting tube method had very high sampling effort, such that the amount of time spent obtaining the readings would, in many cases, be prohibitively high. In contrast, digital photographs are easy and quick to take in the field, and use of automated batch processing decreases the amount of time for which user interaction is necessary in after fieldwork analysis. Total sampling effort is lower than for either the sighting tube method or the moosehorn, and, when the human-based processing time is considered, digital analysis takes the same time as visual estimation when the latter also includes the time taken to transcribe field notes to analysable electronic format—a process undertaken automatically in the digital analysis method—(15 seconds versus 20 seconds, resp.).

*4.2. The Wider Applicability of the Digital Analysis Technique.* The flexible digital image analysis technique described here is suitable for any situation where the purpose is to estimate the percentage cover and/or spatial pattern of one object that can be differentiated clearly from the background matrix in a monochrome picture. Thus it would be suitable for quantifying other aspects of vegetation structural complexity, including percentage ground cover using photographs taken from above and complexity within a vegetation layer (e.g., a scrub layer) using photographs taken horizontally within that layer against a white background (e.g., a sheet).

Time-series photographs could also be used to assess the processes of leaf burst or defoliation. The technique could be used to quantify temporal vegetation colonisation of new surfaces (e.g., new quarry faces or mudflows).

**4.3. Conclusions.** The digital image analysis technique is a quick, inexpensive, and precise method of determining canopy density and dispersion objectively. Tests have shown that it performs to a very high standard in both homogeneous and heterogeneous woodland and that interobserver variability is very low. In comparison with many past attempts to analyse canopy photographs digitally, this method is easier and quicker as spherical geometry is not needed [36]. Moreover, the flexible threshold selection means that there is no requirement to take photographs under uniformly overcast skies as in previous methods [11, 21]. The technique has wider applicability as a tool to survey other aspects of ecological complexity and to quantify temporal or spatial changes in relative abundance of a specific species if that species can be distinguished from its surroundings in a monochrome photograph.

## Acknowledgments

Thanks are due to George Peterken for encouragement and suggesting the need to measure canopy patchiness and Rick Stafford for comments on a draft of this paper. Thanks also are due to the ten volunteers who trialled different methods to enable quantification of interobserver variability: A. Arthur, M. Goodenough, L. Hancock, M. Lee, A. Morss, P. Preston, A. Pyke, S. Pyke, R. Watkins, and T. Watkins.

## References

- [1] C. McElhinny, P. Gibbons, C. Brack, and J. Bauhus, "Forest and woodland stand structural complexity: its definition and measurement," *Forest Ecology and Management*, vol. 218, no. 1–3, pp. 1–24, 2005.
- [2] G. Peterken, *Woodland Conservation and Management*, Chapman and Hall, London, UK, 2nd edition, 1993.
- [3] B. Barnes, D. Zak, S. Denton, and S. Spurr, *Forest Ecology*, John Wiley & Sons, New York, NY, USA, 4th edition, 1998.
- [4] S. R. Englund, J. J. O'Brien, and D. B. Clark, "Evaluation of digital and film hemispherical photography and spherical densiometry for measuring forest light environments," *Canadian Journal of Forest Research*, vol. 30, no. 12, pp. 1999–2005, 2000.
- [5] O. M. Fincke, "Interspecific competition for tree holes: consequences for mating systems and co-existence in neotropical damselflies," *The American Naturalist*, vol. 139, pp. 80–101, 1992.
- [6] J. L. Ganey and W. M. Block, "A comparison of two techniques for measuring canopy density," *Western Journal of Applied Forestry*, vol. 9, pp. 21–23, 1994.
- [7] M. J. Wisdom, R. S. Holthausen, B. C. Wales et al., "Source habitats for terrestrial vertebrates of focus in the interior Columbia basin: broadscale trends and management implications," Tech. Rep. PNW-GTR-485, Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Ore, USA, 2000.
- [8] D. Mudappa, A. Kumar, and R. Chellam, "Abundance and habitat selection of the Malabar spiny dormouse in the rainforests of the southern Western Ghats, India," *Current Science*, vol. 80, no. 3, pp. 424–427, 2001.
- [9] J. L. Ganey, W. M. Block, and S. H. Ackers, "Structural characteristics of forest stands within home ranges of Mexican spotted owls in Arizona and New Mexico," *Western Journal of Applied Forestry*, vol. 18, no. 3, pp. 189–198, 2003.
- [10] D. S. Srivastava, M. C. Melnychuk, and J. T. Ngai, "Landscape variation in the larval density of a bromeliad-dwelling zygopteran *Mecistogaster modesta* (Odonata: Pseudostigmatidae)," *International Journal of Odonatology*, vol. 8, pp. 67–79, 2005.
- [11] G. M. Smith and R. M. Green, "Estimating forest canopy closure using hemispherical photography," *Swansea Geographer*, vol. 31, pp. 1–16, 1994.
- [12] N. Brown, S. Jennings, P. Wheeler, and J. Nabe-Nielsen, "An improved method for the rapid assessment of forest understorey light environments," *Journal of Applied Ecology*, vol. 37, no. 6, pp. 1044–1053, 2000.
- [13] S. Hale, "Managing light to enable natural regeneration in british conifer forests," Forestry Commission Information Note 63, 2004.
- [14] A. C. S. Fiala, S. L. Garman, and A. N. Gray, "Comparison of five canopy cover estimation techniques in the western Oregon Cascades," *Forest Ecology and Management*, vol. 232, no. 1–3, pp. 188–197, 2006.
- [15] S. B. Jennings, N. D. Brown, and D. Sheil, "Assessing forest canopies and understorey illumination: Canopy closure, canopy cover and other measures," *Forestry*, vol. 72, no. 1, pp. 59–73, 1999.
- [16] A. C. Newton, *Forest Ecology and Conservation: A Handbook of Techniques*, Oxford University Press, Oxford, UK, 2007.
- [17] K. J. Kirby, *A Woodland Survey Handbook*, Nature Conservancy Council, Peterborough, UK, 1988.
- [18] M. Kent and P. Coker, *Vegetation Description and Analysis—A Practical Approach*, Wiley Blackwell, London, UK, 1993.
- [19] B. Goldsmith, "Vegetation Monitoring," in *Monitoring for Conservation and Ecology*, F. B. Goldsmith, Ed., pp. 77–86, Chapman and Hall, London, UK, 1990.
- [20] W. J. Sutherland and R. E. Green, "Habitat assessment," in *Bird Ecology and Conservation: A Handbook of Techniques*, W. J. Sutherland, I. Newton, and R. E. Green, Eds., pp. 251–268, Oxford University Press, Oxford, UK, 2004.
- [21] B. M. J. Engelbrecht and H. M. Herz, "Evaluation of different methods to estimate understorey light conditions in tropical forests," *Journal of Tropical Ecology*, vol. 17, no. 2, pp. 207–224, 2001.
- [22] D. A. Clark and D. B. Clark, "Life history diversity of canopy and emergent trees in a neotropical rain forest," *Ecological Monographs*, vol. 62, no. 3, pp. 315–344, 1992.
- [23] P. E. Lemmon, "A spherical densiometer for estimating forest overstorey density," *Forest Science*, vol. 2, pp. 314–320, 1956.
- [24] M. W. Robinson, "An instrument to measure forest crown cover," *Forestry Chronicle*, vol. 23, pp. 222–225, 1947.
- [25] J. G. Cook, T. W. Stutzman, C. W. Bowers, K. A. Brenner, and L. I. Irwin, "Spherical densimeters produce biased estimates of forest canopy cover," *Wildlife Society Bulletin*, vol. 23, no. 4, pp. 711–717, 1995.
- [26] D. J. Vales and F. L. Bunnell, "Comparison of methods for estimating forest overstorey cover. I. Observer effects," *Canadian Journal of Forest Research*, vol. 18, no. 5, pp. 606–609, 1988.

- [27] G. W. Frazer, C. D. Canham, and K. P. Lertzman, "Gap Light Analyzer (GLA), Version 2.0: Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs," Simon Fraser University, British Columbia, and the Institute of Ecosystem Studies, New York, NY, USA, 1999.
- [28] I. Moorthy, J. R. Miller, J. A. J. Berni, P. J. Zarco-Tejada, and L. Qingmou, "Extracting tree crown properties from ground-based scanning laser data," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '07)*, pp. 2830–2832, Barcelona, Spain, 2007.
- [29] K. Yamamoto, K. Kobayashi, T. Nonoda, A. Inoue, and N. Mizoue, "Effect of settings of digital fisheye photography to estimate relative illuminance within forest under low light conditions," *Journal of Forest Research*, vol. 15, no. 5, pp. 283–288, 2010.
- [30] P. J. Radtke and P. V. Bolstad, "Laser point-quadrat sampling for estimating foliage-height profiles in broad-leaved forests," *Canadian Journal of Forest Research*, vol. 31, no. 3, pp. 410–418, 2001.
- [31] S. S. Chan, R. W. McCreight, J. D. Walstad, and T. A. Spies, "Evaluating forest vegetative cover with computerized analysis of fisheye photographs," *Forest Science*, vol. 32, no. 4, pp. 1085–1091, 1986.
- [32] T. C. Whitmore, N. D. Brown, M. D. Swaine, D. Kennedy, C. I. Goodwin-Bailey, and W. K. Gong, "Use of hemispherical photographs in forest ecology: measurement of gap size and radiation totals in a Bornean tropical rain forest," *Journal of Tropical Ecology*, vol. 9, no. 2, pp. 131–151, 1993.
- [33] G. C. Evans and D. E. Coombe, "Hemispherical and woodland canopy photography and the light climate," *Journal of Ecology*, vol. 47, pp. 103–113, 1959.
- [34] P. M. Rich, "Characterizing plant canopies with hemispherical photographs," *Remote Sensing Reviews*, vol. 5, pp. 13–29, 1990.
- [35] P. M. Rich, J. Wood, D. A. Vieglais, K. Burek, and N. Webb, HemiView, Delta-T Devices Ltd, 1999.
- [36] P. L. Mitchell and T. C. Whitmore, *Use of Hemispherical Photographs in Forest Ecology: Calculation of Absolute Amount of Radiation Beneath the Canopy*, Oxford Forestry Institute, Oxford, UK, 1999, Occasional Papers Number 44.
- [37] L. C. Purcell, "Soybean canopy coverage and light interception measurements using digital imagery," *Crop Science*, vol. 40, no. 3, pp. 834–837, 2000.
- [38] M. D. Richardson, D. E. Karcher, and L. C. Purcell, "Quantifying turfgrass cover using digital image analysis," *Crop Science*, vol. 41, no. 6, pp. 1884–1888, 2001.
- [39] M. C. Anderson, "Studies of the woodland light climate. 1. The photographic computation of light conditions," *Journal of Ecology*, vol. 52, pp. 27–41, 1964.
- [40] S. Kato and A. Komiyama, "A calibration method for adjusting hemispherical photographs to appropriate black-and-white images," *The Journal of Forest Research*, vol. 5, pp. 109–111, 2000.
- [41] M. Ishida, "Automatic thresholding for digital hemispherical photography," *Canadian Journal of Forest Research*, vol. 34, no. 11, pp. 2208–2216, 2004.
- [42] M. Nobis and U. Hunziker, "Automatic thresholding for hemispherical canopy-photographs based on edge detection," *Agricultural and Forest Meteorology*, vol. 128, no. 3–4, pp. 243–250, 2005.
- [43] A. Cescatti, "Indirect estimates of canopy gap fraction based on the linear conversion of hemispherical photographs. Methodology and comparison with standard thresholding techniques," *Agricultural and Forest Meteorology*, vol. 143, no. 1–2, pp. 1–12, 2007.
- [44] E. Schwalbe, H. G. Maas, M. Kenter, and S. Wagner, "Hemispherical image modeling and analysis techniques for solar radiation determination in forest ecosystems," *Photogrammetric Engineering and Remote Sensing*, vol. 75, no. 4, pp. 375–384, 2009.
- [45] J. H. Myers, "Selecting a measure of dispersion," *Environmental Entomology*, vol. 7, pp. 619–621, 1978.
- [46] J. Nieder, S. Engwald, M. Klawun, and W. Barthlott, "Spatial distribution of vascular epiphytes (including hemiepiphytes) in a lowland Amazonian rain forest (Surumoni crane plot) of southern Venezuela," *Biotropica*, vol. 32, no. 3, pp. 385–396, 2000.
- [47] F. L. Bunnell and D. J. Vales, "Comparison of methods for estimating forest overstory cover: differences among techniques," *Canadian Journal of Forest Research*, vol. 20, no. 1, pp. 101–107, 1990.
- [48] R. Stafford, A. G. Hart, L. Collins et al., "Eu-social science: the role of internet social networks in the collection of bee biodiversity data," *PLoS ONE*, vol. 5, no. 12, Article ID e14381, pp. 1–7, 2010.