Potential Improvements in the Characterization of Forest Canopy Gaps Caused by Windthrow Using Fine Spatial Resolution Multispectral Data: Comparing Hard and Soft Classification Techniques

Giles M. Foody, Robin G. Jackson, and Christopher P. Quine

**ABSTRACT.** Gaps often form in forest canopies due to windthrow and have important management and ecological implications. Remote sensing has considerable potential for the provision of information on gap properties but this has not been fully realized. This is largely due to the use of conventional (hard, one-pixel one-class) image analysis techniques and imagery with a relatively coarse spatial resolution. This article investigates the potential to extract information on gap properties from fine spatial resolution airborne thematic mapper imagery using soft classification techniques that allow image pixels to have multiple and partial class membership. It is shown that a standard hard maximum likelihood classification may be used to derive an accurate map of the land cover of a forested site (95.1%) from which gaps in a canopy of Sitka spruce were accurately identified (94.5%). The maximum likelihood classification was also softened by outputting probabilities of class membership for each pixel. Softening the classification increased the information on gap properties that could be extracted from the data. In particular, the accuracy with which key gap properties, such as gap area, perimeter length and shape, were estimated was higher in the outputs of the softened than hard classification. Thus, while strong correlations between the remotely sensed and ground data estimates of gap area ($r \geq 0.96$) and perimeter ($r \geq 0.87$), based on a sample of 36 gaps, were derived from all classifications, the accuracy with which gap properties were estimated was generally highest when a soft classification was used. For example, the use of a soft rather than hard classification resulted in the root mean square error in estimating gap area declining from 144.90 to 132.87 m². Furthermore, the soft classification allowed the sharpness of the gap boundary to be estimated, enabling further gap properties to be inferred. In particular, the soft classification output enabled the direction of the wind event causing the initial damage to be estimated, and it may aid the definition of sites with a future risk of windthrow. *For. Sci.* 49(3):444–454.

**Key Words:** Gap boundary, soft classification.
Many forests are vulnerable to wind damage. In the United Kingdom, for example, forests have often been planted on shallow or poorly drained soils. As a result, rooting may be limited and the trees sensitive to windthrow. The risk of windthrow constrains many silvicultural practices (Quine and Miller 1990, Quine et al. 1995). Decisions on which species to plant and whether or not to thin a stand, for example, are typically taken with reference to the potential for windthrow. However, strong winds, particularly those associated with the passage of Atlantic depression systems, have led to the formation of gaps by windthrow in the canopies of many British forests (Miller 1985, Quine and Bell 1998). Although these gaps can be viewed positively as, for example, a means of increasing structural and ecological diversity (Quine et al. 1999), windthrow is typically perceived as undesirable from an economic standpoint, especially if the forest is managed for wood production (Coutts 1986).

Although gaps caused by windthrow are a common component of forests, our understanding of their formation and later dynamics as well as their effects on airflow over the forest and potential for future wind damage is poor (Gray and Spies 1996 Quine and Bell 1998). This is due, in part, to the difficulty of acquiring information on gaps at appropriate spatial and temporal scales. Field survey of gaps can be detailed, but sometimes access is impossible or the site unsafe for survey personnel. Moreover, the repeated coverage of large areas is both difficult and costly. Remote sensing offers the potential to study gaps caused by windthrow at a variety of spatial and temporal scales. Furthermore, it has the potential to provide complete coverage of large regions, irrespective of safety or access problems. Although the spatial resolution of popular satellite sensor systems may be too coarse for many forest management applications, airborne sensors and those carried on a suite of new satellites are able to provide fine spatial resolution data (<5 m) (Aplin et al. 1997). These data may be well suited to the provision of information at the forest-landscape scale (200–2,000 ha), which is that used frequently by forest managers in Europe (Quine and Bell 1998). In particular, remote sensing may be used to identify gaps and characterize their boundaries. In addition to basic inventory applications, information on gap properties may be useful for planning applications, particularly in regard to assessing the potential for future windthrow (Gardiner and Quine 2000). In this way, remote sensing may help in the development of silvicultural prescriptions to maximize the potential return from managed forests and reduce the windthrow risk associated with certain management operations.

While the increase in availability of fine spatial resolution data should aid forest management applications, some problems remain. One key concern relates to the reliance on conventional “hard” image analysis techniques to extract information from the remotely sensed data (Foody 1996). These techniques are termed “hard” since each pixel is associated fully with a single class, irrespective of how appropriate this association may be. Thus, these techniques assume that each pixel represents a homogeneous region of a single class. This assumption of pure pixels is, even in fine spatial resolution data, often untenable (Campbell 1996, p. 273–277). This article aims to evaluate the potential to identify and characterize gaps caused by windthrow from fine spatial resolution imagery using “soft” image analysis techniques. These techniques allow multiple and partial class membership, which are characteristic features of mixed pixels that are commonly encountered in remote sensing, as well as providing a representation that may be more appropriate for classes that vary continuously (Foody 1996, 2000). In particular, the potential to derive thematic maps that portray the spatial distribution of forest land cover classes and characterize gap properties (e.g., area, perimeter, shape, and boundary sharpness) is assessed.

**Thematic Mapping**

Mapping land cover from digital remotely sensed data is often achieved through the application of a supervised image classification (Mather 1999, p. 174–188). This has three main stages. First is the training stage in which the spectral response of sites of known class membership in the image is statistically characterized. The training statistics derived are used in the second, class allocation, stage of the classification to allocate each pixel (other defined unit) to one of the predefined classes. This allocation is made on the basis of relative similarity of the pixel to each class, with the allocated class being the one to which the pixel has greatest similarity. Thus, for example, in the widely used maximum likelihood classification, each pixel is allocated to the class with which it has the highest likelihood of membership (Foody et al. 1992, Mather 1999, p. 181–185). The third and final stage of the classification is the testing stage in which the accuracy of the classification is assessed. Therefore, the final product is effectively a thematic map with a specified level of accuracy.

Supervised classification is one of the most commonly encountered analyses in remote sensing. It is commonly used in forestry applications, especially in studies focusing on mapping and monitoring forests with particular regard to changes in forest cover and forest succession (e.g., Skole and Tucker 1993, Foody et al. 1996, Steininger 1996, Helmer et al. 2000) in addition to its use as an input to ecosystem simulation models that include land cover as a driving variable (e.g., Lucas and Curran 1999). There are, however, many problems with supervised classification. Often the accuracy of the classification is insufficient for operational use (Wilkinson 1996) and the resulting thematic map is a fundamentally flawed representation of land cover (Woodcock and Gopal 2000). As each pixel is allocated to a single class by a hard classification, a major problem is the representation of discrete objects such as gaps in a forest canopy. A gap will be represented in the output of a conventional hard classification as a “blocky” region with the gap-canopy boundary constrained to lie between image pixels (Atkinson 1997, Foody 1998). This boundary will also be of constant sharpness along the full length of the gap’s perimeter. This representation is unfortunate as the pixel is an arbitrary
Softened Image Classification

A soft or fuzzy classification does not force each pixel to belong to a single class but instead allows each pixel to have multiple and partial class memberships (Wang 1990, Foody 1996). This type of classification can be derived in a variety of ways but especially through the softening of a conventional hard image classification. Thus, for example, the maximum likelihood classification may be softened to output measures of the strength of class membership rather than just the code of the most likely class of membership (Foody et al. 1992). Two measures of class membership are of particular interest. First, the posterior probability (likelihood) of a pixel belonging to a specified class. This indicates the uncertainty of a particular class allocation and is made relative to the membership to all defined classes. Second, the typicality of class membership. This is measured with regard to a single class of interest and represents how typical a case is of the selected class. Both of these measures can be expressed as probabilities and are effectively derived in the determination of class membership, though discarded, in a conventional hard classification. Further information on the derivation of these metrics is given in Foody et al. (1992). Relative to the conventional hard classification, the information content of the remotely sensed data may be more fully exploited by outputting these probabilities to form a soft classification. Preliminary work has indicated that a soft classification highlights gap properties, such as the location of exposed rootplates and variations in boundary sharpness (Jackson et al. 2000). In this article, the analyses are extended to focus on the properties of the gap boundary in greater detail.

Test Site

The test site was the Cwm Berwyn forest in central Wales, UK. This forest was established over three planting seasons between 1960 and 1963 by the Forestry Commission and is typical of upland forest in Britain. Our attention focused particularly on the part of the forest established as a wind damage monitoring site in 1988 by the Forestry Commission (Quine and Reynard 1990). This comprised seven compartments that contained many gaps formed by windthrow and surrounded a small lake, Llyn Berwyn (Figure 1). The windthrow monitoring site was planted mainly with stands of Sitka spruce (Picea sitchensis [Bong.] Carr.) with smaller areas planted with species such as Japanese larch (Larix kaempferi [Lamb.] Carr.).

The Cwm Berwyn forest is located in an area of moderately variable topography with elevation ranging between 300 and 500 m above sea level. The site is exposed with a TOPEX ranging between 1 and 34 (Miller 1985, Quine and Bell 1998). TOPEX is an index of topographic exposure obtained by summing the skyline angles taken to the visible horizon at the eight principal points of the compass (Miller et al. 1987, Quine and White 1998). TOPEX has been found to have a strong association with wind exposure (Miller et al. 1987), and the larger the TOPEX value, the greater the degree of topographic shelter.

The soils at the site are typically peats and gleys, and as the annual rainfall is high (approximately 1,700 mm yr$^{-1}$), the site was drained by ploughing prior to planting. Even so, the rooting of trees at the site will be shallow, and the plough furrows will restrict root spread, leading to enhanced risk of windthrow. The threat of windthrow may be summarized by the windthrow hazard classification (Miller 1985). This combines information related to the windiness of a site (e.g., topographic exposure, elevation) with that on tree anchorage (e.g., soil type) to derive a score on a 1–6 scale to indicate the risk of windthrow. A low score indicates that a site that is unlikely to suffer wind damage whereas a high score high-

Figure 1. Test site and data. (a) ATM image, band 10 of the site and (b) map of the gaps in the forest at the test site identified with confidence from the aerial photographs.
lights a substantial risk of windthrow (Miller 1985, Quine and Bell 1998). The windthrow hazard classification for the test site varies in score from 3 to 6 (Miller 1985), with a mean of 4.4 for the whole forest (Quine and Bell 1998), indicating a significant risk of windthrow.

Within the wind damage monitoring area, no restrictions were placed on stand thinning, although two compartments of Sitka spruce were left as unthinned controls. In addition, no clearance of wind damaged trees or removal of windthrown trees along newly exposed edges was permitted during the monitoring period (Quine et al. 1995). As windthrown trees were not removed, fallen trees generally dominated the interior of each gap.

Data and Methods

The formation of gaps within the forest canopy due to the action of windthrow and their subsequent change over time have been monitored at the test site. This was achieved mainly through the manual interpretation of color aerial photographs acquired annually since 1988 (except in 1993) at a 1:10,000 scale. The aerial photographs were used to identify and map gaps in the forest canopy. The gaps mapped included those identified with a very high degree of confidence as well as those that may possibly be gaps but for which an identification could not be made with complete certainty. The interpreted boundaries of the gaps identified from the aerial photographs were digitized and their coordinates, in terms of the UK national projection, stored within a geographical information system (GIS). This information on gap properties is supported by observations from field surveys undertaken throughout the monitoring site that acquired detailed data on stand and site conditions.

Remotely sensed data of the site were acquired in July 1994 with a Daedalus 1268 airborne thematic mapper (ATM) sensor (Wilson 1997). The ATM is an electro-optical line scanning sensor similar to systems carried on Landsat and other satellites. It acquired data in 11 spectral wavebands sampled from the blue (0.42–0.45 μm) to thermal infrared (8.50–13.00 μm) regions of the electromagnetic spectrum, with 7 wavebands broadly similar to those used by the Landsat TM (Wilson 1997). These ATM data were acquired from an altitude of approximately 2,000 m above the ground and as a consequence have a spatial resolution of approximately 4 m. This spatial resolution is finer than that derived from popular civilian satellite systems such as the SPOT HRV (10 m panchromatic or 20 m multispectral) and Landsat TM (30 m in nonthermal wavebands and 120 m in the thermal waveband) but similar to the spatial resolution of a suite of sensors that have been launched recently or are scheduled for launch in the near future (Aplin et al. 1997). The ATM data were preprocessed to reduce image distortions. Although a radiometric correction was not required as the classifications are based on relative spectral differences (Mather 1999, p. 67) the ten nonthermal wavebands were calibrated to radiance using the sensor’s gain and offset coefficients; the information required to rigorously calibrate the data acquired in the thermal waveband were unavailable. The data set was also geometrically registered to the UK national grid, and thereby to the data on forest properties held within the GIS, using a second-order polynomial equation (Figure 1). This transformation equation was derived using 49 ground control points dispersed across the ATM image. To preserve the statistical properties of the ATM data, nearest neighbor resampling was used, and the estimated root mean square error of the geometric transformation was 1.76 pixels. This was larger than the usual target, typically <0.5 pixel, but as precise overlay of the data sets was unnecessary, it was not considered detrimental to the research. In this study, the geometrical transformation was used simply to aid the location of specific gaps on the ground and in the remotely sensed data sets. Since these data sets were analyzed separately, precise coregistration was unnecessary, although this would be desirable in an operational application.

The wind damage monitoring site comprised five thematic classes. These were the moorland surrounding the forest, water in Llyn Berwyn, the gaps in the forest canopy formed by windthrow and the unbroken forest canopies of Sitka spruce and Japanese larch. Training sites to describe each of the five classes for the classification analyses were identified in the ATM data. A total of 808 pixels were used to train the classifications. This sample was comprised of 192 pixels of Sitka spruce forest canopy, 111 pixels of Japanese larch canopy, 194 pixels of moorland, 174 pixels of water, and 137 pixels of gap. These training pixels were all separated by a minimum distance of 10 pixels to reduce the deleterious effects of spatial autocorrelation on the derived training statistics (Curran and Atkinson 1998). The sample for each class was also defined to satisfy, as far as possible, the recommended sample size suggested for use with probabilistic classifiers. Typically it is recommended that the training sample for each class be at least 10–30 times the number of wavebands used in the classification (Piper 1992, Mather 1999, p. 175).

A feature selection analysis was undertaken to reduce the data volume without significant loss of discriminatory information (Mather 1999, p. 203–206). Here, a stepwise discriminant analysis was used to determine an appropriate waveband combination to use in the analyses of the ATM data. On the basis of the discriminant analysis, four wavebands, numbers 4 (visible red, 0.60–0.63 μm), 7 (near-infrared, 0.76–0.90 μm), 9 (shortwave infrared, 1.55–1.75 μm) and 11 (thermal infrared, 8.50–13.00 μm), which represented the main dimensions of the ATM data set, were used in the research.

The training statistics derived from the ATM data in the four selected wavebands were used to produce a conventional (hard) maximum likelihood classification of the ATM data. This hard classification was also softened to provide outputs that depicted for each pixel the posterior and typicality probabilities of class membership. The ability to identify and characterize gaps in the classification outputs was evaluated relative to the ground data on gap properties, especially those derived from the interpretation of the 1994 aerial photography held within the GIS.

Gaps formed prior to the ATM data acquisition in July 1994 and identified from the aerial photography were the
main focus of the research. At the time of the ATM data collection, gaps caused by windthrow ranged in size from 41 to 2,158 m² with a median area of 152.5 m² (Figure 2), typical of the size of gaps observed in British forests (Quine and Bell 1998). The identification of gaps required only an evaluation of their presence or absence in the ATM and ground data sets. The ability to characterize gap properties such as the area, perimeter, shape, and boundary sharpness, however, required more detailed ground data. The data on the gap boundaries in the GIS were used to estimate gap area, perimeter and shape. These properties were estimated for 36 gaps of variable size distributed across the windthrow monitoring site. This sample represented all the gaps identified by the aerial photograph interpretation with a high degree of confidence that were spatially isolated from other noncanopy features (e.g., other gaps, unplanted strips between compartments) by at least 3 pixels or 12 m. This enabled the work to focus on only gap properties. It also removed the need for precise coregistration of the ATM and ground data sets, because once the gap had been identified it could be characterized in each data set separately.

The area and shape of the gaps was estimated from the boundary depicted in the image classifications. The area of each gap was estimated from the hard classification by summing the area of adjacent pixels classified as belonging to the gap class. With the soft classifications, the area contained within specified probability of class membership contours fitted to the classification output was used to represent gap area. A range of probability contours were evaluated, but on the basis of earlier work (Foody et al. 1999, Jackson et al. 2000) attention here focuses particularly on the 0.5 posterior probability of membership to forest and the 0.79 and 0.84 typicality of membership to gap probability contours as means of representing the gap boundary (Figure 3). Once defined, the gap boundary was also the basis of further characterizations of the gap. The length of the boundary, for example, provided an estimate of gap perimeter. The derived boundary of each gap was also used to characterize gap shape. Shape may be quantified in a variety of ways and was expressed here as the fractal dimension, calculated as twice the slope of the log-log regression of gap perimeter versus gap area [which has been used in other studies to quantify landscape components from remotely sensed data (Wickham et al. 1997) and in studies of gaps (Quine and Bell 1998)].

The boundary of a gap is typically not of uniform sharpness along its perimeter. The sharpness of the boundary generally varies systematically in relation to the direction of the wind event that initiated its formation. The sharpest and most gradual parts of the boundary correspond to the locations where windthrown trees have fallen away from the boundary and where the fallen trees lean into standing trees respectively (Figure 4). Ground data on the sharpness of the gap boundaries were acquired in two ways. First, a field survey of the 36 selected gaps was undertaken in July 1999. Since this was ~5 yr after the acquisition of the ATM data, and it was unsafe to survey the gaps in detail, this field survey aimed to identify only the general characteristics of the gap boundaries. In particular, the survey aimed to identify the sharpest and most gradual parts of the boundary.

In recognition of the problems with this method, notably the 5 yr time interval between the field survey and the ATM data acquisition, a second data set on gap boundary sharpness was acquired. This was based on a digital elevation model (DEM) of the Cwm Berwyn forest derived by digital photogrammetric techniques from color aerial photography acquired in 1995, approximately 1 yr after the ATM data acquisition (Miller et al. 1997). The DEM had a horizontal spatial resolution of 1 m and a vertical accuracy (RMS) of ±0.7 m; further details on the DEM are provided in Miller et al. (2000). The DEM was coregistered to the classification outputs to enable the sharpness of the gap boundary in each to be compared. For this, a map depicting the gradient in change in canopy height (in degrees) was derived from the DEM by linear interpolation. As attention was only focused on the general nature of the variations in sharpness around the gap and the location of the sharpest and most gradual locations, this map depicting the rate of change in canopy height, was classified to show quartiles of boundary sharpness. These two reference data sets on gap boundary sharpness, derived from the field survey and DEM, were compared to the outputs of the image classifications. This comparison focused on the soft classification outputs as the gap boundary is constrained to be of uniform sharpness in the output of the hard classification.

The sharpness of the gap boundary was mapped from the soft classification outputs. This was achieved by estimating the gradient in class membership strength between two class membership contours in the soft classification outputs. These were the 0.84 and 0.89 typicality and 0.5 and 0.9 posterior probability contours. The gradient of class membership estimated from each of the soft classifications was taken to be positively related to the sharpness of the boundary. To simplify the investigation, the boundary of the gap depicted in the soft classification outputs was divided into four quartiles of sharpness (Figure 3f). The compass bearing from the midpoint of the region primarily representing the sharpest to that of the most gradual region of the boundary, derived manually, followed the direction of tree fall and hence of the wind event that caused the gap to form. This direction of tree...
fall was also measured in the field for 34 of the 36 gaps; for 2 gaps, the pattern of treefall was too complex to unambiguously identify treefall direction in the field.

**Results and Discussion**

The conventional hard classification of the ATM data resulted in the production of a five-class thematic map of the site. The accuracy of this classification was assessed using a sample of 1,000 pixels that were independent of the pixels used to derive the training statistics. Overall 95.1% of these testing pixels were correctly allocated (Table 1). In terms of identifying land cover, the conventional hard classification appeared to provide an accurate representation of the site.

Although the hard classification appeared to be accurate, this does not necessarily mean it appropriately identifies and represents the gaps in the forest canopy. Focusing on the 183 gaps identified by aerial photograph interpretation that lie within the Sitka spruce forest canopy, the ability of the hard classification to identify and represent gaps was evaluated. Of the 183 gaps present, 173 (94.5%) were correctly identified by the hard classification. It was also apparent that the classification depicted other parts of the site as belonging to

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese larch</td>
<td>19 2 0 4 0 25 76.00</td>
</tr>
<tr>
<td>Sitka spruce</td>
<td>431 9 0 0 441 97.73</td>
</tr>
<tr>
<td>Gap</td>
<td>1 14 145 7 0 167 86.82</td>
</tr>
<tr>
<td>Moorland</td>
<td>1 3 5 281 1 291 96.56</td>
</tr>
<tr>
<td>Water</td>
<td>0 0 1 75 76 98.68</td>
</tr>
<tr>
<td>Σ</td>
<td>22 450 159 293 76 1000</td>
</tr>
<tr>
<td>Producer’s accuracy (%)</td>
<td>86.36 95.77 91.19 95.90 98.68</td>
</tr>
</tbody>
</table>
the gap class. This appears in Table 1 as an apparent commission error or over-estimation of the gap class. However, closer inspection revealed that most of these instances related to sites where trees were widely spaced due to drain lines or, more importantly, corresponded to sites identified by the aerial photograph interpreters as possible or potential gaps. The hard classification of the ATM data may, therefore, be identifying some gaps that had not been identified with confidence from the aerial photographs. A further 32 gaps were known to exist in the area planted with Japanese larch and all (100%) of these were identified correctly by the hard classification. These gaps, however, were generally complex and often merged. As a result of this and because the Japanese larch class was the least accurately classified (Table 1) and only occupied a small portion of the test site, regions of this class were excluded from further analysis.

The area, perimeter, and shape of the 36 isolated gaps in the Sitka spruce forest were estimated from the hard classification of the ATM data. Although the area estimated from the classification was strongly correlated with the reference data \( r = 0.96 \), the root mean square error (RMSE) was relatively large (144.90 \( \text{m}^2 \)) (Figure 5). Similarly, the estimated perimeter of the gaps derived from the ATM data was strongly correlated with the reference data \( r = 0.87 \) but the RMSE of the estimation was again large (40.68 m) (Figure 6). As the gaps were relatively small features, the proportion of pixels at or near the boundary was large, and hence the effect of misallocation around the boundary was likely to be significant and responsible for some of the error. In the estimation of both the area and perimeter of the gaps, a major problem was the blocky nature of the output from the hard classification, a consequence of each pixel being allocated fully to a single class (Figure 3). This forced the gap boundary to be located between pixels in the output of the classification.

Since the real gap boundary was unlikely to ever truly lie between pixels but rather lie within them, the softened classification output may be expected to provide a more appropriate and accurate representation of the gaps. Estimates of gap area derived from the softened classification were at least as strongly correlated \( (r \geq 0.96) \) with the reference data as those from the hard classification (Figure 5). However, the calculated RMSE of the estimates were still large (up to 170.31 \( \text{m}^2 \)). Similarly, the gap perimeters were generally estimated slightly more accurately from the softened rather than hard classification output (Figure 6). This was most evident when the gaps were represented by the 0.84
typicality probability contour fitted to the soft classification output. In this instance, the correlation between the predicted and reference data on gap perimeter was 0.90 (for the hard classification $r = 0.87$) and the RMSE of the estimates was 37.53 m (for the hard classification RMSE = 40.68 m). These results indicate that soft classifications may be used to derive a slightly more accurate characterization of gap area and perimeter than conventional hard classifications. It is likely that different trends would be observed at different spatial resolutions but at the 4 m spatial resolution of the ATM data set used, the difference in the estimates derived from the hard and soft classifications was small (e.g., for estimates of gap perimeter, the difference in correlation was no more than 0.03 and the largest difference in the RMSE was 7.42 m).

The softened classifications did, however, appear to provide a visually more realistic representation of the gaps and also provided a richer description of the gaps and their boundaries in particular. The estimated shape of the gaps, expressed as the fractal dimension, was more accurate from the softened than hard classifications. The RMSE in the estimates of the fractal dimension, for example, declined from 0.128 for the hard classification to 0.012, 0.085, and

Figure 5. Relationships between the area of gaps predicted from the image classification output and that contained in the reference data set. (a) Hard classification ($r = 0.96; \text{RMSE} = 144.90 \text{ m}^2$), (b) Soft classification with the gap boundary represented by the 0.5 posterior probability of membership to forest contour ($r = 0.96; \text{RMSE} = 146.05 \text{ m}^2$), (c) Soft classification with the gap boundary represented by the 0.94 typicality probability of membership to gap ($r = 0.97; \text{RMSE} = 132.87 \text{ m}^2$), (d) Soft classification with the gap boundary represented by the 0.79 typicality probability of membership to gap ($r = 0.97; \text{RMSE} = 170.31 \text{ m}^2$). Note, natural logarithms were used in the production of the figure.
Figure 6. Relationships between the perimeter of gaps predicted from the image classification output and that contained in the reference data set: (a) Hard classification \((r = 0.87; \text{RMSE} = 40.68 \text{ m})\), (b) Soft classification with the gap boundary represented by the 0.5 posterior probability of membership to forest contour \((r = 0.87; \text{RMSE} = 48.10 \text{ m})\), (c) Soft classification with the gap boundary represented by the 0.84 typicality probability of membership to gap \((r = 0.90; \text{RMSE} = 37.53 \text{ m})\), (d) Soft classification with the gap boundary represented by the 0.79 typicality probability of membership to gap \((r = 0.88; \text{RMSE} = 38.09 \text{ m})\). Note, natural logarithms were used in the production of the figure.

0.002 for the estimates derived from the boundaries represented by the class membership contours fitted at 0.5 posterior probability, 0.79, and 0.84 typicality probabilities respectively. More importantly, the soft classification could also be used to quantify the variation in the sharpness of the gap boundary along its entire length; with the hard classification sharpness is constant along the length of the gap’s perimeter.

The sharpness of the gaps depicted in the soft classifications was compared to that in the two reference data sets. To aid the comparison with the ground data sets, the data on boundary sharpness derived from the soft classifications were reduced to quartiles. Focusing on the quartiles representing the sharpest and most gradual portions of the boundary for each gap, it was apparent that the location of these regions corresponded with the maps derived from the field survey. This was confirmed for pixels in both the sharpest and most gradual quartiles with chi-square tests comparing the quartile memberships of the boundary pixels in the two data sets (significant at the 99% level of confidence). Comparison of the boundary sharpness derived from the soft classifications with the rate of change in canopy height derived from the DEM revealed the same trend. Although this analysis could only be undertaken for the 10 largest gaps in the
sample, due to problems in relating the DEM to the ATM data, the quartiles representing the extreme rates of change in the DEM and soft classification were positively associated (chi-square statistics significant at the 99% level of confidence). The soft classification output, therefore, appeared to provide information on the variation in boundary sharpness. Moreover, this appeared to allow the direction of tree fall and hence of the wind event causing the initiation of the gap to be inferred. Comparison of the estimated treefall direction derived from the soft classification with the treefall direction measured in the field showed a high degree of correspondence. This was confirmed with a Hotelling’s test for paired sample angles (Zar 1999, p. 461–462), which revealed no significant difference in the angle of treefall between the two data sets (at the 95% level of confidence).

The ability to characterize gap shape and map the sharpness of the gap boundary with a softened classification may help predict sites of future risk of windthrow. The most gradual boundary, for instance, can gently direct the airflow over the boundary reducing the wind loading on the trees that form the gap-canopy boundary (Quine et al. 1995, Gardiner and Stacey 1996). This, therefore, provides some degree of wind protection unlike the sharp boundary where the boundary edge trees are more vulnerable to damage due to the removal of neighboring trees (Quine et al. 1995). The precise effect of variations in the sharpness of the boundary varies with the shape of the gap relative to the wind direction which can act to determine the size of the area affected (Quine and Miller 1990, Kimmins 1997, p. 224–226). Thus the potential of soft classifications to provide information on gap shape and sharpness may make them an attractive tool to help further the understanding of windthrown gap properties, although there are substantial challenges to implementing them operationally.

**Conclusions**

Fine spatial resolution remotely sensed data and soft image classification techniques appear to have considerable potential for the derivation of useful information on forest canopy gaps. In particular, the results demonstrate that a conventional hard classification may be softened and a variety of gap properties such as area, perimeter, shape, and boundary sharpness extracted from the classification outputs.

While the hard classification provided an accurate representation of the land cover at the site and correctly identified most gaps, its representation of the gaps was limited. In particular, the blocky nature of the hard classification output provided a visually unrealistic representation of the gaps. Relative to the hard classification from which it was derived, the softened classification outputs represented the gaps more appropriately and allowed more accurate estimation of the gap properties. Perhaps the biggest and most important difference between the hard and soft classification outputs was the ability to characterize variations in boundary sharpness from the soft classification. The information on boundary sharpness also enabled the direction of treefall to be accurately inferred from the soft classification outputs. This ability to describe the boundary in detail, together with more accurate estimates of gap area and shape, combine to make the soft classification output more useful for describing the existing gaps and also for assessing the implications of the gaps for future windthrow. Fine spatial resolution remotely sensed data and soft image classification techniques may, therefore, substantially aid the understanding and prediction of gap formation and progression.

**Literature Cited**


