Agricultural and Forest Meteorology 247 (2017) 181–193

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Validating canopy clumping retrieval methods using hemispherical photography in a simulated *Eucalypt* forest

William Woodgate[a,b,c,*], John D. Armston[d,e,f], Mathias Disney[g,h], Lola Suarez[a,b], Simon D. Jones[a,b], Michael J. Hill[i,j], Phil Wilkes[a,b,k], Mariela Soto-Berelov[a,b]

[a] School of Mathematical and Geospatial Sciences, RMIT University, GPO Box 2476V, Melbourne, VIC 3001, Australia
[b] Cooperative Research Centre for Spatial Information, Carlton, VIC 3053, Australia
[c] Land and Water, Commonwealth Scientific and Industrial Research Organisation, Yarralumla, ACT, Australia
[d] Remote Sensing Centre, Science Delivery, Queensland Department of Science, Information Technology, and Innovation, 41 Boggo Road, QLD 4102, Australia
[e] Joint Remote Sensing Research Program, School of Geography, Planning and Environmental Management, University of Queensland, St. Lucia, QLD 4072, Australia
[f] Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA
[g] Department of Geography, University College London,Gower Street, London WC1E 6BT, UK
[h] NERC National Centre for Earth Observation, UK
[i] Department of Earth System Science and Policy, University of North Dakota, Clifford Hall, 9011, 4149 University Drive, Grand Forks, ND 58202, USA
[j] School of the Environment, Flinders University, Sturt Road, Bedford Park, South Australia 5042, Australia
[k] ITC, University of Twente, P.O. Box 217, NL-7000 AE Enschede, The Netherlands

ARTICLE INFO

Keywords:
Leaf Area Index
Clumping
Woody-to-total area
3D modelling
Hemispherical photography
Eucalypt

ABSTRACT

The so-called clumping factor (Ω) quantifies deviation from a random 3D distribution of material in a vegetation canopy and therefore characterises the spatial distribution of gaps within a canopy. Ω is essential to convert *effective* Plant or Leaf Area Index into *actual* LAI or PAI, which has previously been shown to have a significant impact on biophysical parameter retrieval using optical remote sensing techniques in forests, woodlands, and savannas. Here, a simulation framework was applied to assess the performance of existing *in situ* clumping retrieval methods in a 3D virtual forest canopy, which has a high degree of architectural realism. The virtual canopy was reconstructed using empirical data from a Box Ironbark Eucalypt forest in Eastern Australia. Hemispherical photography (HP) was assessed due to its ubiquity for indirect LAI and structure retrieval. Angular clumping retrieval method performance was evaluated using a range of structural configurations based on varying stem distribution and LAI. The CLX clumping retrieval method (Leblanc et al., 2005) with a segment size of 15° was the best performing clumping method, matching the reference values to within 0.05 Ω on average near zenith. Clumping error increased linearly with zenith angle to > 0.3 Ω (equivalent to a 30% PAI error) at 75° for all structural configurations. At larger zenith angles, PAI errors were found to be around 25–30% on average when derived from the 55–60° zenith angle. Therefore, careful consideration of zenith angle range utilised from HP is recommended. We suggest that plot or site clumping factors should be accompanied by the zenith angle used to derive them from gap size and gap size distribution methods. Furthermore, larger errors and biases were found for HPs captured within 1 m of unrepresentative large tree stems, so these situations should be avoided in practice if possible.

1. Introduction

The fluxes of radiation, heat and water in a vegetation canopy are primarily determined by the total amount of vegetation and spatial distribution, characterised by Plant and Leaf Area Index or 'PAI' and 'LAI' respectively (Law et al., 2001; Spanner et al., 1990). LAI is recognised as an essential climate variable and key input into global climate models among other applications (GCOS, 2011). It is usually defined as one half the total green leaf area per unit horizontal ground surface area (GCOS, 2011).

LAI is typically estimated indirectly *in situ* from optical remote sensing instruments, which measure the proportion and spatial distribution of gaps in plant canopies (see reviews by (Bréda, 2003; Jonckheere et al., 2004; Zheng and Moskal, 2009)). These methods often utilise a theoretical model of canopy gap probability, *Pgap*, to estimate LAI (Monsi and Saeki, 1965; Nilson, 1971; Woodgate et al.,

* Corresponding author at: Room 27, Level 11, Building 12, Remote Sensing Centre, School for Mathematical and Geospatial Sciences, RMIT University, Melbourne, Australia.
E-mail address: william.woodgate@rmit.edu.au (W. Woodgate).

http://dx.doi.org/10.1016/j.agrformet.2017.07.027
Received 26 February 2017; Received in revised form 3 July 2017; Accepted 31 July 2017
Available online 06 October 2017

2015a). The spatial distribution is characterised by the angular distribution of material in the canopy and the so-called clumping index or factor 'Ω', which describes the typically non-random (clumped) distribution of material in a canopy over a defined 3D volume (Chen and Black, 1991, 1992; Nilson, 1971).

Clumping factors can vary with spatial scale and viewing angle, and are typically in the range of 0.4–1 at the stand scale for broadleaf forests (Chen et al., 2005; Leblanc et al., 2005; Zhao et al., 2012). Clumping factors < 1 imply that canopy elements are arranged in such a way that more gaps are found than if they were distributed randomly. Conversely, clumping factors > 1 are associated with regularly distributed canopy elements, meaning a lower gap proportion than if they were distributed randomly. A clumping factor of 1 is associated with a theoretically random spatial distribution of canopy elements. A main assumption of the *Pgap* model is that the clumping factor is equal to 1 unless explicitly accounted for, and therefore the model provides an estimate of effective (non-clumped) PAI or LAI (Chen and Cihlar, 1995a). However, in forests a clumping factor of 1 is typically invalid due to the clumping of canopy elements at all scales, from shoot/twig, to branch, to crown, to whole canopy or stand (Chen and Black, 1991; Fournier et al., 1997; Pisek et al., 2011). This can lead to up to a 50% LAI underestimation if the clumping factor is 0.5 for example (Leblanc et al., 2005).

In situ clumping estimates are used to support calibration and validation activities of Earth Observation global LAI and clumping products (Fernandes et al., 2014; Pisek et al., 2015). *In situ* clumping retrieval methods are an area of ongoing development and research (see Section 2). The traditional approach to benchmark and evaluate retrieval method performance has been to compare with other *in situ* methods deemed to be most accurate, such as destructive harvesting, allometric relationships and litter traps (c.f. Leblanc et al., 2014). However, this approach is no longer viable when trying to achieve a high level of accuracy in the order of < 10% LAI uncertainty, especially given errors also arise in the benchmark *in situ* methods in forests and woody ecosystems where clumping is particularly important. For example, Chen et al. (1997) reported difficulty in keeping total LAI error budgets below 25% using the direct and semi-direct techniques of harvesting and extrapolation to larger areas via allometry. Similar error budgets have been reported for litter-trap collection methods (Kalácska et al., 2005). This high uncertainty tolerance is unacceptable for current LAI product calibration and validation accuracy targets, which aim to match within 20% of independently derived *in situ* estimates, with the requirement to improve that to within 5% for future applications (Fernandes et al., 2014).

Three-dimensional modelling and computer simulation frameworks are an attractive alternative to traditional field-based benchmarking approaches, e.g. Leblanc and Fournier (2014); Walter et al. (2003); Widlowski et al. (2015); Woodgate et al. (2016). These frameworks enable retrieval method accuracy to be quantified precisely using the known model 'truth' as a benchmark for comparison. Of the few framework studies that assess LAI and clumping retrieval methods, ray tracing models coupled with a limited degree of canopy architectural realism were typically employed. For example, Jonckheere et al. (2006) simulated Hemispherical Photography (HP) in beech and Scots pine forest canopies to evaluate the divergence of theoretical foliage distribution models to the reference model value. Walter et al. (2003) evaluated two angular clumping retrieval techniques using simulated HPs. Virtual scenes comprised flat foliage (10 cm × 10 cm × 1 cm dimensions) of varying densities and clustering percentage. Gonsamo and Pellikka (2009) utilised the virtual scenes from Walter et al. (2003) to investigate five clumping retrieval methods from HPs on level and sloped terrain. More recently, Leblanc and Fournier (2014) evaluated the accuracy of four indirect clumping retrieval methods. HPs were simulated in virtual forest scenes comprising a broad range of stand structures.

A main challenge of modelling frameworks is the adequate representation of canopy architectural realism, where the spatial distribution of gaps at all scales should closely reflect empirical data. These simulation studies cited above found retrieval methods need to be further tested under conditions where canopies better represent the structural properties of an actual forest stand, such as including internal branching structure. The advantage in this case is to permit better definition of the accuracy, strengths and limitations of these indirect estimation methods. Consequently, there is a need for more complex 3D models created specifically for this purpose that accurately reflect the tree and stand structure of real forests.

The primary objective of this study was to determine the accuracy of several clumping retrieval methods when applied to a mature Eucalypt woodland. HP was assessed due to its ubiquitous and low-cost use for indirect LAI and structural retrieval (Leblanc and Fournier, 2014). Secondary objectives included quantitatively determining how the accuracy of the clumping retrieval methods vary with view zenith angle, stem distribution, and stand LAI. A simulation framework was applied to a virtual forest canopy with a high degree of architectural realism reconstructed from empirical data, representative of a Box Ironbark Eucalypt forest in Eastern Australia (Woodgate et al., 2016).

This manuscript first reviews briefly the theoretical background of *in situ* clumping retrieval methods (Section 2). Next, the 3D modelling and simulation framework is presented (Section 3) enabling the evaluation of clumping method performance from HP's to a range of stem distributions and LAI values (Section 4). The paper concludes with a discussion on the practical implications and recommendations for users of *in situ* clumping retrieval methods (Sections 5 and 6).

2. Theoretical background: *in situ* clumping retrieval methods

Canopy clumping retrieval methods are used to determine the degree of non-randomness of an observed canopy. The following sections present a categorisation and description of *in situ* clumping retrieval methods, typically applied from measurements of gap size and gap size distribution. As such, the clumping factor is a function of view zenith angle and can be retrieved over narrow angle ranges from instrument measurements such as HP images (Leblanc and Chen, 2001; Leblanc and Fournier, 2014; Pisek et al., 2011).

2.1. Logarithmic averaging (Ω_{LX})

Lang and Xiang (1986) proposed a method to retrieve clumping based on logarithmic averaging, using finite segments '*k*' of gap size distribution measurements, hereafter referred to as LX or Ω_{LX}:

$$\Omega_{LX}(\theta) = \ln[\overline{Pgap(\theta)}]/\overline{\ln[Pgap(\theta)]} = n\ln[\overline{Pgap(\theta)}]/\sum_{k=1}^{n}\ln[Pgap_k(\theta)] \quad (1)$$

Where $\ln[\overline{Pgap(\theta)}]$ is the logarithm of averaged gap fraction over a predefined area (i.e. a proportion of a single measurement of gap sizes or multiple measurements), $\overline{\ln[Pgap(\theta)]}$ is the average logarithm of gap fraction over the same area. $Pgap_k(\theta)$ is the gap fraction of segment k relating to a sub-domain of the gap size measurement. For example, $k = 90^\circ$ equates to the 360° azimuth ring of a HP image being divided into four 90° segments containing gap and canopy pixels. The size of k should preferably be at least 10 times the mean canopy element width, typically corresponding to leaf size. Two assumptions are made: (i) the canopy elements at the k scale are distributed randomly, and (ii) segments contain gaps, due to the undefined logarithm of $Pgap_k = 0$.

Due to the undefined calculation of a logarithm with a null-gap segment, a gap value of half a pixel is assigned. The maximum LAI value within a segment therefore becomes a function of k and zenith angle. Gonsamo et al. (2010) proposed a clumping retrieval method based on the LX principle, that utilises the minimum segment size for which a gap is present, as opposed to a fixed segment size that could lead to null