Citation
Hingee, K. and Baddeley, A. and Caccetta, P. and Nair, G. 2019. Computation of Lacunarity from Covariance of Spatial Binary Maps. Journal of Agricultural, Biological, and Environmental Statistics. 24 (2): pp. 264-288. http://doi.org/10.1007/s13253-019-00351-9

Supplementary Material

October 5, 2018

Contained in this supplementary document is:

Appendix A	A table giving formulae for the gliding box lacunarity (GBL) of various transformations of stationary RACS.
Appendix B	An explicit relation between $\widehat{\mathrm{L}}_{\mathrm{GB}}$ and the traditional covariance estimator.
Appendix C	A brief report on Mandelbrot's $\mathcal{L}_{M1}$ index and its relation to GBL and contact distributions.
Appendix D	Detailed methods, results and analysis for the simulation study summarised in Section 5.
Appendix E	Additional information, including the data and R code, for our application of GBL estimators to a time series of meso-scale forest maps in Section 6.1.
Appendix F	Further results for the application to decimetre resolution tree canopy maps in Section 6.2.
Appendix G	A description of the asymptotic computational cost of the GBL estimators.

Other supplementary material is `stationaryracsinference_0.4-01.tar.gz`, which contains an R package with the functions and related tools for computing every GBL estimator that we investigated, and the following files related to our demonstration of GBL estimators on meso-scale forest maps (see Appendix E for details): `satelliteimages`, `finalcloudandshadowmasks.RData`, `finalmaskedforests.RData`, `manualforrepeatinganalysis.Rnw`, `gbltrads.RData`, `gblcs.RData`, `gblccs.RData`, and `gblgs.RData`.

Contents

A GBL of Intersections, Unions and Invertible Linear Transformations

Pattern	Coverage Probability	Covariance	GBL
$T(\mathbb{X}_1)$	p_1	$C_1(T^{-1}(\mathbf{v}))$	$\frac{1}{p_1^2\|B\|^2}\int \gamma_B(\mathbf{v})C_1(T^{-1}(\mathbf{v}))d\mathbf{v}$
$\mathbb{X}_1 \cap \mathbb{X}_2$	$p_1 p_2$	$C_1(\mathbf{v})C_2(\mathbf{v})$	$\frac{1}{p_1^2 p_2^2\|B\|^2}\int \gamma_B(\mathbf{v})C_1(\mathbf{v})C_2(\mathbf{v})d\mathbf{v}$
$\mathbb{X}_1 \cup \mathbb{X}_2$	$p_1 + p_2 - p_1 p_2$	$C_1(\mathbf{v}) + C_2(\mathbf{v}) + 2p_1p_2 - 2C_1(\mathbf{v})p_2 - 2C_2(\mathbf{v})p_1 + C_1(\mathbf{v})C_2(\mathbf{v})$	$\frac{1}{(p_1+p_2-p_1p_2)^2\|B\|^2}\int \gamma_B(\mathbf{v})\Big(C_1(\mathbf{v}) + C_2(\mathbf{v}) + 2p_1p_2 - 2C_1(\mathbf{v})p_2 - 2C_2(\mathbf{v})p_1 + C_1(\mathbf{v})C_2(\mathbf{v})\Big)d\mathbf{v}$

Table 1: The coverage probability, covariance and GBL of RACS derived from an invertible linear transformation T with inverse T^{-1}, intersection or union of independent stationary RACS, $\mathbb{X}_1$ and $\mathbb{X}_2$. The RACS, $\mathbb{X}_1$ and $\mathbb{X}_2$, have coverage probability p_1, p_2 and covariance C_1, C_2 respectively.

B The Relation Between $\widehat{\mathrm{L}}_{\mathbf{GB}}$ and Covariance

The exact relation between $\widehat{\mathrm{L}}_{\mathrm{GB}}(B)$ and $\hat{C}(\mathbf{v})$ can be obtained using reasoning similar to that used to prove (33) and might be used to bound the difference between $\widehat{\mathrm{L}}_{\mathrm{GB}}(B)$ and $\widehat{\mathrm{L}}_{C}(B)$ with a function of $\hat{p}$ and $\hat{C}(\mathbf{v})$.

Theorem 3 *Suppose that a set X with positive volume is observed in an observation window W and that the region of interest, Z, is taken to be W. Then the moments (15) and (16) used in $\widehat{\mathrm{L}}_{GB}$ are*

$$\frac{1}{|W \ominus \breve{B}|}\int_{W\ominus\breve{B}} |X \cap (B \oplus \{\mathbf{y}\})| d\mathbf{y} = \hat{p}|B|\frac{|W|}{|W \ominus \breve{B}|} - \frac{1}{|W \ominus \breve{B}|}\int_{X\cap W} f(\mathbf{x})d\mathbf{x} \quad \text{(B.1)}$$

and

$$\begin{aligned}\frac{1}{|W \ominus \breve{B}|}\int_{W\ominus\breve{B}} |X \cap (B \oplus \{\mathbf{y}\})|^2 d\mathbf{y} = &\int_{\mathbb{R}^d} \frac{\gamma_W(\mathbf{z})}{|W \ominus \breve{B}|}\gamma_B(\mathbf{z})\hat{C}(\mathbf{z})d\mathbf{z} \\ &- \frac{1}{|W \ominus \breve{B}|}\int_{X\cap W}\int_{\mathbb{R}^d} g(\mathbf{x},\mathbf{z})\mathbf{1}_{X\cap W}(\mathbf{z}+\mathbf{x})d\mathbf{z}d\mathbf{x},\end{aligned} \quad \text{(B.2)}$$

where $f(\mathbf{x}) = |B| - |(W \ominus \breve{B}) \cap (\breve{B} \oplus \{\mathbf{x}\})|$ *and* $g(\mathbf{x}, \mathbf{z}) = \gamma_B(\mathbf{z}) - |(W \ominus \breve{B}) \cap (\breve{B} \oplus \{\mathbf{x}\}) \cap (\breve{B} \oplus \{\mathbf{z} + \mathbf{x}\})|$, *which are such that*

$$x \in W \ominus (\breve{B} \oplus B) \implies f(\mathbf{x}) = 0 \text{ and } g(\mathbf{x}, \mathbf{z}) = 0, \text{ and} \tag{B.3}$$

$$x + \mathbf{z} \in W \ominus (\breve{B} \oplus B) \implies g(\mathbf{x}, \mathbf{z}) = 0. \tag{B.4}$$

Proof We start with the proof of (B.1). Recall that the mass $|X \cap (B \oplus \{\mathbf{y}\})|$ can be written as an integral of indicator functions

$$|X \cap (B \oplus \{\mathbf{y}\})| = \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_{B \oplus \{\mathbf{y}\}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) d\mathbf{x}.$$

Thus the first moment (15) is

$$\begin{aligned} \frac{1}{|W \ominus \breve{B}|} \int_{W \ominus \breve{B}} |X \cap (B \oplus \{\mathbf{y}\})| d\mathbf{y} &= \frac{1}{|W \ominus \breve{B}|} \int_{W \ominus \breve{B}} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \frac{1}{|W \ominus \breve{B}|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{W \ominus \breve{B}}(\mathbf{y}) \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \frac{1}{|W \ominus \breve{B}|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_{W \ominus \breve{B}}(\mathbf{y}) \mathbf{1}_{\breve{B} \oplus \{\mathbf{x}\}}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \frac{1}{|W \ominus \breve{B}|} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) |(W \ominus \breve{B}) \cap (\breve{B} \oplus \{\mathbf{x}\})| d\mathbf{x} \end{aligned}$$

If we define the function $f(\mathbf{x})$ as

$$f(\mathbf{x}) := |B| - |(W \ominus \breve{B}) \cap (\breve{B} \oplus \{\mathbf{x}\})|$$

then

$$\frac{1}{|W \ominus \check{B}|} \int_{W \ominus \check{B}} |X \cap (B \oplus \{\mathbf{y}\})| d\mathbf{y} = \frac{1}{|W \ominus \check{B}|} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x})(|B| - f(\mathbf{x})) d\mathbf{x}$$
$$= \frac{|B||X \cap W|}{|W \ominus \check{B}|} - \frac{1}{|W \ominus \check{B}|} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$
$$= \hat{p}|B| \frac{|W|}{|W \ominus \check{B}|} - \frac{1}{|W \ominus \check{B}|} \int_{X \cap W} f(\mathbf{x}) d\mathbf{x}.$$

46 This proves (B.1).

The volume $|(W \ominus \check{B}) \cap (\check{B} \oplus \{\mathbf{x}\})|$ is the volume of the set of box centres, $\mathbf{y}$, such that the corresponding box, $B \oplus \{\mathbf{y}\}$, both contains $\mathbf{x}$ and is completely contained in W. When $\mathbf{x}$ is away from the edge of W then this volume is simply $|B|$ and $f(\mathbf{x}) = 0$,

$$\begin{aligned}
\mathbf{x} \in W \ominus (\check{B} \oplus B) &\implies \mathbf{x} \oplus B \oplus \check{B} \subseteq W \\
&\implies \check{B} \oplus \{\mathbf{x}\} \subseteq W \ominus \check{B} \\
&\implies |(W \ominus \check{B}) \cap (\check{B} \oplus \{\mathbf{x}\})| = |\check{B} \oplus \{\mathbf{x}\}| = |B| \\
&\implies f(\mathbf{x}) = 0.
\end{aligned}$$

47 This proves the required property for $f(\mathbf{x})$.

Using similar techniques the second moment (16) is

$$\begin{aligned}
&\frac{1}{|W \ominus \check{B}|} \int_{W \ominus \check{B}} |X \cap (B \oplus \{\mathbf{y}\})|^2 d\mathbf{y} \\
&= \frac{1}{|W \ominus \check{B}|} \int_{W \ominus \check{B}} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) d\mathbf{x} \int_{\mathbb{R}^d} \mathbf{1}_{X \cap W}(\mathbf{z}) \mathbf{1}_B(\mathbf{z} - \mathbf{y}) d\mathbf{z} d\mathbf{y} \\
&= \frac{1}{|W \ominus \check{B}|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{W \ominus \check{B}}(\mathbf{y}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) \mathbf{1}_B(\mathbf{z} - \mathbf{y}) \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_{X \cap W}(\mathbf{z}) d\mathbf{x} d\mathbf{z} d\mathbf{y} \\
&= \frac{1}{|W \ominus \check{B}|} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{W \ominus \check{B}}(\mathbf{y}) \mathbf{1}_B(\mathbf{x} - \mathbf{y}) \mathbf{1}_B(\mathbf{z} + \mathbf{x} - \mathbf{y}) \mathbf{1}_{X \cap W}(\mathbf{x}) \mathbf{1}_{X \cap W}(\mathbf{z} + \mathbf{x}) d\mathbf{x} d\mathbf{z} d\mathbf{y}.
\end{aligned} \tag{B.5}$$

Note that

$$\begin{aligned}\int_{\mathbb{R}^d} \mathbf{1}_{W\ominus\check{B}}(\mathbf{y})\mathbf{1}_B(\mathbf{x}-\mathbf{y})\mathbf{1}_B(\mathbf{z}+\mathbf{x}-\mathbf{y})d\mathbf{y} &= \int_{\mathbb{R}^d} \mathbf{1}_{W\ominus\check{B}}(\mathbf{y})\mathbf{1}_{\check{B}\oplus\{\mathbf{x}\}}(\mathbf{y})\mathbf{1}_{\check{B}\oplus\{\mathbf{z}+\mathbf{x}\}}(\mathbf{y})d\mathbf{y} \\ &= |(W\ominus\check{B})\cap(\check{B}\oplus\{\mathbf{x}\})\cap(\check{B}\oplus\{\mathbf{z}+\mathbf{x}\})|.\end{aligned}$$

We will denote the difference between the above and $\gamma_B(\mathbf{z})$ as

$$g(\mathbf{x},\mathbf{z}) = \gamma_B(\mathbf{z}) - |(W\ominus\check{B})\cap(\check{B}\oplus\{\mathbf{x}\})\cap(\check{B}\oplus\{\mathbf{z}+\mathbf{x}\})|.$$

The function $g(\mathbf{x},\mathbf{z})$ is zero if either $\mathbf{x}$ or $\mathbf{z}+\mathbf{x}$ are in $W\ominus(\check{B}\oplus B)$ as $\mathbf{x}\in W\ominus(\check{B}\oplus B)$ implies that $(W\ominus\check{B})\cap(\check{B}\oplus\{\mathbf{x}\}) = \check{B}\oplus\{\mathbf{x}\}$ and

$$\begin{aligned}|(W\ominus\check{B})\cap(\check{B}\oplus\{\mathbf{x}\})\cap(\check{B}\oplus\{\mathbf{z}+\mathbf{x}\})| &= |(\check{B}\oplus\{\mathbf{x}\})\cap(\check{B}\oplus\{\mathbf{z}+\mathbf{x}\})| \\ &= |(\check{B}\oplus\{\mathbf{z}\})\cap\check{B}| \\ &= \gamma_B(\mathbf{z}),\end{aligned}$$

and similarly for $\mathbf{z}+\mathbf{x}$.

Continuing from (B.5) we get

$$\begin{aligned}&\frac{1}{|W\ominus\check{B}|}\int_{W\ominus\check{B}} |X\cap(B\oplus\{\mathbf{y}\})|^2 d\mathbf{y} \\ &= \frac{1}{|W\ominus\check{B}|}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\gamma_B(\mathbf{z})-g(\mathbf{x},\mathbf{z}))\mathbf{1}_{X\cap W}(\mathbf{x})\mathbf{1}_{X\cap W}(\mathbf{z}+\mathbf{x})d\mathbf{x}d\mathbf{z} \\ &= \frac{1}{|W\ominus\check{B}|}\left(\int_{\mathbb{R}^d}\gamma_B(\mathbf{z})\gamma_{X\cap W}(\mathbf{z})d\mathbf{z} - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(\mathbf{x},\mathbf{z})\mathbf{1}_{X\cap W}(\mathbf{x})\mathbf{1}_{X\cap W}(\mathbf{z}+\mathbf{x})d\mathbf{x}d\mathbf{z}\right) \\ &= \int_{\mathbb{R}^d}\frac{\gamma_W(\mathbf{z})}{|W\ominus\check{B}|}\gamma_B(\mathbf{z})\hat{C}(\mathbf{z})d\mathbf{z} - \frac{1}{|W\ominus\check{B}|}\int_{X\cap W}\int_{\mathbb{R}^d} g(\mathbf{x},\mathbf{z})\mathbf{1}_{X\cap W}(\mathbf{z}+\mathbf{x})d\mathbf{z}d\mathbf{x}.\end{aligned}$$

This is the result claimed in (B.2).

C Mandelbrot's $\mathcal{L}_{M1}$ Index

Here we relate $\mathcal{L}_{M1}$, which we defined in (18), to GBL and spatial statistics' contact distributions (Section C.3). We then compare estimates of $\mathcal{L}_{M1}$ for a binary map of forest locations to results from the FracLac package for ImageJ (Karperien, 2015). Our estimates agreed up to an offset of 1 with results from one of the methods in FracLac and suggest that users of FracLac, such as Dàvila and Parés (2007), have used $\mathcal{L}_{M1}$.

In the following we give additional background (Section C.1), and introduce contact distributions of stationary RACS (Section C.2). We then derive the relation between $\mathcal{L}_{M1}$, GBL and contact distributions (Section C.3), and compare $\mathcal{L}_{M1}$ estimates to FracLac results (Section C.4).

C.1 Background

Mandelbrot (1983, p315) in a section called 'lacunarity as second-order effect concerning the mass prefactor' described a possible lacunarity index as $\mathbb{E}\left[(M/\mathbb{E}[M]-1)^2\right]$ where $\mathbb{E}[M]$ was the expected mass of a random fractal within a fixed region, assuming that the random fractal intersected the fixed region. If we replace the random fractal with a stationary RACS that produces topologically regular[1] closed sets then we get

$$\mathcal{L}_{M1}(B) := \frac{\mathrm{Var}\big(\,|B\cap\mathbb{X}|\,\big|\,|B\cap\mathbb{X}|>0\big)}{\mathbb{E}\big[\,|B\cap\mathbb{X}|\,\big|\,|B\cap\mathbb{X}|>0\big]^2}, \tag{C.1}$$

where B denotes a fixed region.

The manual for the FracLac package (Karperien, 2005, p26) describes an estimator of a lacunarity index as using all boxes that are tested and an estimator of another lacunarity index as 'counting only boxes having pixels'. Since the former represents all possible box locations, the latter must be a subset of the box locations and it seems likely that the latter uses only boxes that contain *foreground* pixels, which would correspond to an estimator of $\mathcal{L}_{M1}$.

FracLac has been used by a number of authors, including Dàvila and Parés (2007) (and Dàvila et al. (2007)), who used a lacunarity index to study plasma protein gels. Dàvila and Parés suggest that the lacunarity index that they used was closely related to a coefficient of variation, however the index was not GBL as Figure 5 of (Dàvila and Parés, 2007) contains estimates that approach zero for small boxes whilst estimates of

[1] A subset A of $\mathbb{R}^d$ is called topological regular if it is equal to the closure of its interior, $A=\overline{A^o}$.

GBL approach $1/\hat{p}$ for small boxes, where $\hat{p}$ is the coverage probability estimate. It seems likely that the index estimated by Dàvila and Parés was $\mathcal{L}_{M1}$.

C.2 Spatial Statistics' Contact Distribution

Given a convex set B containing the origin, the unconditional contact distribution of stationary RACS, $\mathbb{X}$, is defined as (Hansen et al., 1999)

$$F_B^u(r) := \begin{cases} P(\mathbb{X} \cap (\mathbf{x} \oplus rB) \neq \varnothing) & \text{if } r \geqslant 0 \\ 0 & \text{if } r < 0, \end{cases} \tag{C.2}$$

where B is called the *gauge body*, $rB = \{r\mathbf{x} : \mathbf{x} \in B\}$ is the gauge body scaled by r, $\varnothing$ is the empty set, and the location $x \in \mathbb{R}^d$ is arbitrary due the stationarity of $\mathbb{X}$. Estimators of $F_B^u(r)$ are described by Chiu et al. (2013, §6.4.5) and are available in the spatstat package in R (Baddeley et al., 2015).

C.3 Relation to GBL and Contact Distributions

The following gives a relation between $\mathcal{L}_{M1}(rB)$, $\mathrm{L}(rB)$ and $F_B^u(r)$. As contact distributions are infinite order properties of RACS the relation shows that $\mathcal{L}_{M1}$ is also an infinite order property. The usual estimators of $F_B^u(r)$ are closely related to box-counting dimension estimators[2] and so from a purely empirical perspective estimates of $\mathcal{L}_{M1}$ are confounded with both GBL estimates and box-counting dimension estimates.

Theorem 4 *Suppose $\mathbb{X}$ is a stationary RACS with positive coverage probability and that B is a convex set with positive volume such that $o \in B$, then*

$$\mathcal{L}_{M1}(rB) = \mathrm{L}(rB) F_B^u(r) - 1. \tag{C.3}$$

[2] this seems to be rarely stated and will be discussed elsewhere

Proof Let $Y = |B \cap \mathbb{X}|$. Then

$$\mathbb{E}[Y] = P(Y > 0)\mathbb{E}[Y|Y > 0] + 0 \tag{C.4}$$

and

$$\mathbb{E}[Y^2] = P(Y > 0)\mathbb{E}[Y^2|Y > 0]. \tag{C.5}$$

The following completes the proof

$$\begin{aligned}
\mathcal{L}_{M1}(rB) &= \frac{\mathrm{Var}\big(\,|B \cap \mathbb{X}|\,\big|\,|B \cap \mathbb{X}| > 0\big)}{\mathbb{E}\big[\,|B \cap \mathbb{X}|\,\big|\,|B \cap \mathbb{X}| > 0\big]^2} && \text{(C.6)}\\
&= \frac{\mathrm{Var}(Y|Y > 0)}{\mathbb{E}[Y|Y > 0]^2} && \text{(C.7)}\\
&= \frac{\mathbb{E}[Y^2|Y > 0] - \mathbb{E}[Y|Y > 0]^2}{\mathbb{E}[Y]^2/P(Y > 0)^2} && \text{(C.8)}\\
&= \frac{\frac{\mathbb{E}[Y^2]}{P(Y>0)} - \left(\frac{\mathbb{E}[Y]}{P(Y>0)}\right)^2}{\mathbb{E}[Y]^2/P(Y > 0)^2} && \text{(C.9)}\\
&= \frac{\mathbb{E}[Y^2]P(Y > 0) - \mathbb{E}[Y]^2}{\mathbb{E}[Y]^2} && \text{(C.10)}\\
&= \frac{\mathbb{E}[Y^2]}{\mathbb{E}[Y]^2}P(Y > 0) - 1 && \text{(C.11)}\\
&= \mathrm{L}(rB)F_B^u(r) - 1. && \text{(C.12)}
\end{aligned}$$

C.4 Comparison to FracLac

We estimated GBL given square boxes from the binary map in Figure 1 using $\widehat{\mathrm{L}}_{\mathrm{GB}}$ and also estimated $F_B^u(r)$ for a square gauge body using a Kaplan-Meier estimator (Hansen et al., 1999). These estimates, which we denote $\widehat{\mathrm{L}}_{\mathrm{GB}}(rB)$ and $\hat{F}_B^u(r)$ respectively, were plugged into (C.3) to estimate $\mathcal{L}_{M1}$,

$$\hat{\mathcal{L}}_K(rB) = \hat{F}_B^u(r)\,\widehat{\mathrm{L}}_{\mathrm{GB}}(rB) - 1. \tag{C.13}$$

To the same map we applied FracLac's sliding box (SLAC) methods (Karperien, 2005, p25-26) for lacunarity indices (Karperien, 2015). For this computation the parameters in

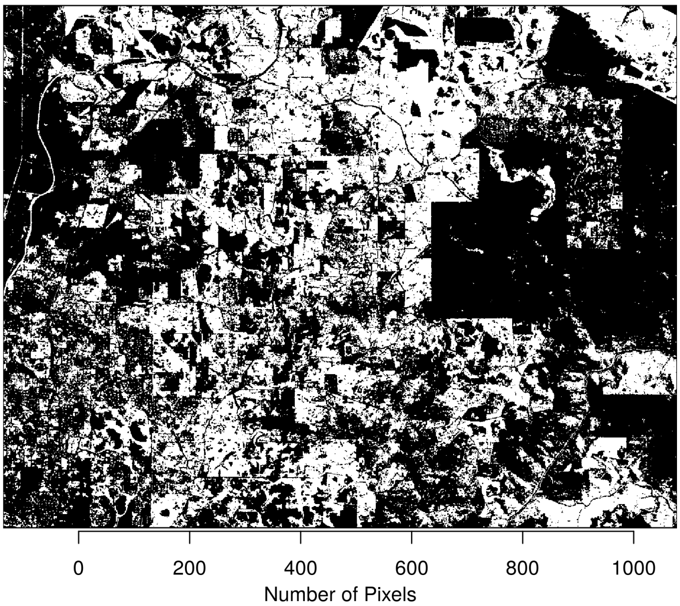

Figure 1: A presence-absence map of forest. FracLac's 'SLAC' computations and estimators of $\mathcal{L}_{M1}$, GBL and $F^u_B(r)$ were applied to this map. Black: Forest. White: Non-forest.

FracLac were chosen such that the box location shifted 1 pixel width at a time and 58 box sizes were used, ranging from 2 pixels to 25% of the image size.

The estimated $\hat{\mathcal{L}}_K(rB)$, $\hat{F}^u_B(r)$ and $\widehat{\mathrm{L}}_{\mathrm{GB}}(rB)$ are shown in Figure 2 with some of the results of the FracLac computations. Note that the estimated $\hat{F}^u_B(r)$ contains steps due to the discrete pixel size in the binary map which is a likely cause of the saw-like features in $\hat{\mathcal{L}}_K(rB)$. The lacunarity indices reported by FracLac with columns titled '$(\sigma/\mu)^2 + 1$ for $L\Omega$' and '$(\sigma/\mu)^2 + 1$ for $F(mass)$' matched our $\widehat{\mathrm{L}}_{\mathrm{GB}}(rB)$ and $\hat{\mathcal{L}}_K(rB) + 1$, respectively. Furthermore the estimates $\hat{\mathcal{L}}_K(rB)$ here have a similar form to Figure 5 of (Dàvila and Parés, 2007) and support our suspicion that Dàvila and Parés used $\mathcal{L}_{M1}(rB)$ estimates.

D Simulation Study of GBL Estimators

In this section we use simulations of a stationary RACS to compare the performance of GBL estimators under different scenarios.

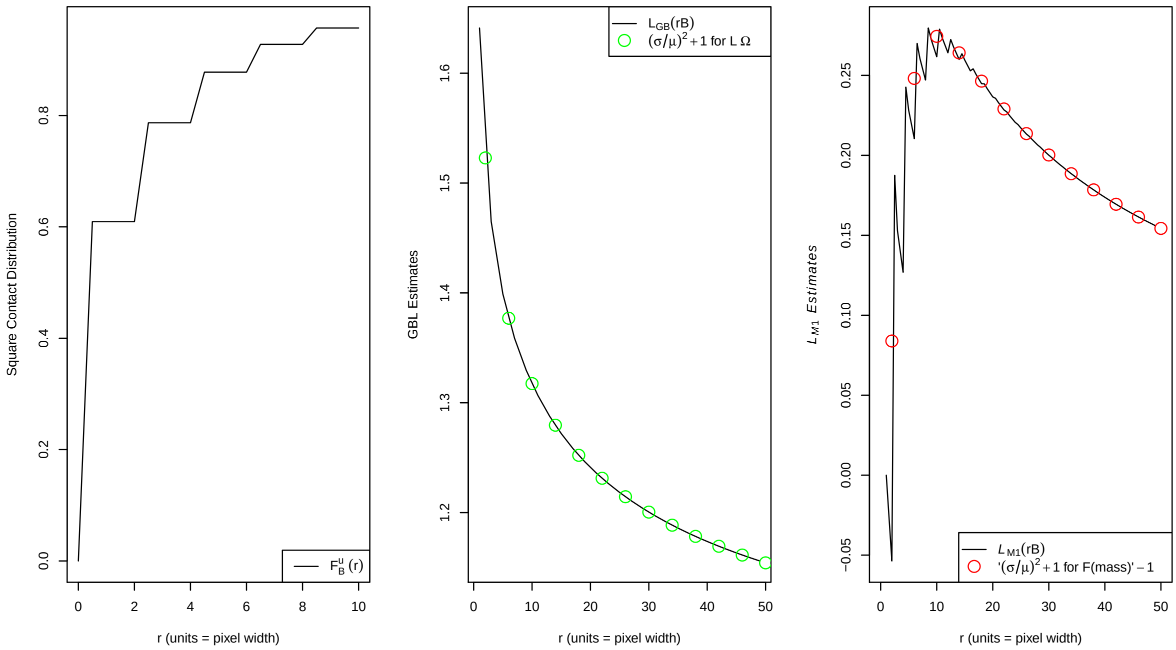

Figure 2: Our estimates (solid lines) of $F^u_B(r)$ (left), $\mathrm{L}(rB)$ (centre) and $\mathcal{L}_{M1}(rB)$ (right) with results from FracLac (circles). Here B is a square of unit width centred on the origin; further details in main body of text. *In centre:* FracLac results are from the column titled '$(\sigma/\mu)^2+1$ for $L\Omega$'. *In right:* FracLac results are from the column titled '$(\sigma/\mu)^2+1$ for $F(mass)$' and offset by 1.

D.1 Methods

We used simulated raster binary maps of a 200 unit × 200 units study region, denoted Z. The pixels in these binary maps were 0.1 units wide. The foreground was created by discretising realisations of a stationary RACS, $\mathbb{X}$, to the pixel grid. In many cases we applied a pattern of occlusions so that the study region Z was not fully observed. If we denote a pattern of occlusions by $A \subset \mathbb{R}^d$, discretised to the pixel grid, then in these cases the observation window, W, of the binary map was $W = Z \backslash A$.

The stationary RACS $\mathbb{X}$ that we simulated for the foreground of the binary maps was a Boolean model[3] with 0.005 expected germs per unit area and grains that were discs centred on the origin. The radius of the discs was distributed according to the discrete approximation, at integer units of radius, of the probability density function,

$$f(r) = \begin{cases} 0 \text{ if } r < 1, \\ \frac{k}{r^2} \text{ if } 1 \leqslant r \leqslant 50, \\ 0 \text{ if } r > 50, \end{cases} \tag{D.1}$$

where r is the radius and k is a normalising constant. This distribution for the radius was similar to the size distribution of discs in Mandelbrot's disc tremas (Mandelbrot, 1983, §33) and was chosen so that $\mathbb{X}$ exhibited some multiscale behaviour. The coverage probability of $\mathbb{X}$ was about $p = 0.44$. The covariance of $\mathbb{X}$ can be calculated using the set covariance of the discs (Chiu et al., 2013, eq. 1.58, 3.18) and the pair-correlation of $\mathbb{X}$ (Figure 3) was such that the probability of points being covered by $\mathbb{X}$ was independent for points further than 100 units from each other. Note that the width of the study region, Z, was twice this distance. The RACS $\mathbb{X}$ was simulated using the function `rbpto` in the attached R package. Due to time restriction we do not investigate foreground simulated by other RACS.

We considered three scenarios for which GBL estimators might be used:

- *Scenario 1:* Realisations of $\mathbb{X}$ were observed in fixed windows. These observation

[3] A *Boolean model* is a stationary RACS that is a union of identically distributed independent random sets (called *grains*) centred on the points (called *germs*) of a stationary Poisson point process (Chiu et al., 2013, §3).

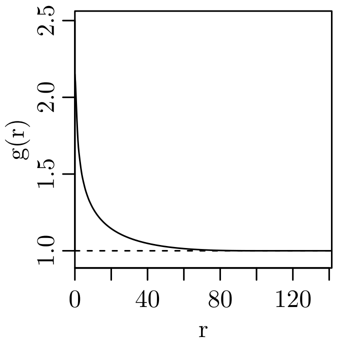

Figure 3: The pair-correlation $g(r)$ of the foreground process, $\mathbb{X}$, for points separated by a distance r.

windows were the full study region Z, Z excluding given various patterns of occlusions, and a square 40 units wide. The occlusion patterns used covered 2%, 31%, 50%, 70%, and 90% of the study region. The 40 units wide observation window was included to investigate the performance of the estimators when the observation window is much smaller than the spatial interaction distance of the RACS observed. For each observation window 1000 realisations of $\mathbb{X}$ were simulated.

The occlusion pattern that covered 2% of the study region was a realisation of a Boolean model with grains that were discs of radius 2.5 and germ intensity of 0.001. The occlusion patterns that covered 31%, 50%, 70%, and 90% of Z were realisations of Boolean models with a germ intensity of 0.005 germs per unit area and grains that were discs of deterministic radius equal to $5, 6.8, 8.8$, and 12, respectively. The observation windows given by these occlusions, along with the full study region and the 40 units wide observation window are shown with example realisations of $\mathbb{X}$ in Figure 4.

- *Scenario 2:* A realisation, X, of $\mathbb{X}$ was fixed and observed in the study region Z excluding random patterns of occlusions. The occlusion patterns were generated according to a Boolean model, $\mathbb{O}$, that had a germ intensity of 0.001 germs per unit area, grains that were discs with radius equal to 5, and a coverage probability of about 0.076. For this scenario 1000 patterns of occlusions were simulated and the realised coverage fraction of the occlusions ranged from 4.4% to 12% of Z.

- *Scenario 3:* A collection of binary maps with foreground given by $\mathbb{X}$ and observation window given by $Z \backslash \mathbb{O}$ were simulated, where $\mathbb{O}$ was the same Boolean model for occlusions used in Scenario 2. This scenario thus combined the sources of variability in Scenario 1 and Scenario 2. Here $\mathbb{X}$ and the occlusion pattern $\mathbb{O}$ were simulated 1000 times.

The estimators $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{C}$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$, $\widehat{\mathrm{L}}_{\kappa M}$, $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$, defined in (14) and (25) - (30), were applied to the simulated binary maps using functions in the attached R package. Note that, following conventional procedure, $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was applied by replacing the region of interest in (15) and (16) with the observation window.

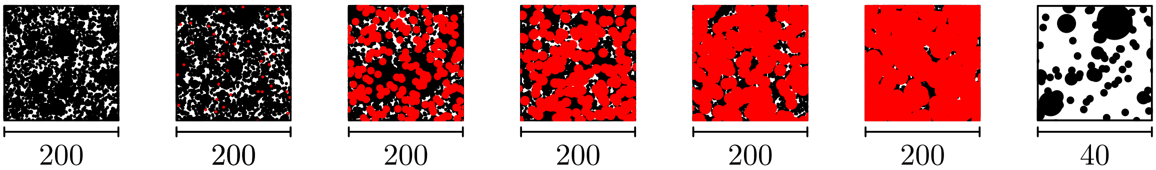

Figure 4: The observation windows used in *Scenario 1* with example simulations of $\mathbb{X}$. *Red:* Occlusions - these were not part of the observation window. *Black:* Foreground. *White:* Background. *From left:* Study region without occlusions, the patterns of occlusion that covered 2%, 31%, 50%, 70% and 90% of the study region, the square observation window 40 units wide.

D.2 Results and Analysis

In the following results for each scenario a discussed individually. For Scenario 1 and Scenario 2 (22) enables the GBL of $\mathbb{X}$ to be computed from the covariance of $\mathbb{X}$ and allows us to assess the bias of the GBL estimators.

D.2.1 Scenario 1

The pointwise mean and pointwise variance of the estimators using the fixed observation windows of Scenario 1 are shown in Figure 5. Figure 6 shows the pointwise bias of the estimators and the pointwise variance relative to the pointwise variance of $\widehat{\mathrm{L}}_{\kappa H}$.

Fully Observed Study Region When the study region was fully observed all estimators had similar small bias. The variance of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ and the balanced covariance-based estimators, $\widehat{\mathrm{L}}_{gM}, \widehat{\mathrm{L}}_{gI}, \widehat{\mathrm{L}}_{\kappa M}, \widehat{\mathrm{L}}_{\kappa I}$, and $\widehat{\mathrm{L}}_{\kappa H}$, were very similar to each other, and were substantially smaller than the variance of the $\widehat{\mathrm{L}}_{C}$ estimator for box widths greater than 15.

The larger variance observed for $\widehat{\mathrm{L}}_C$ seems likely to be related to correlations between the numerator and denominator in each estimator; Figure 7 shows that $\widehat{\mathrm{L}}_C$ given boxes wider than 15 had smaller Pearson correlation between estimated variance of $|B \cap \mathbb{X}|$ (numerator) and estimated mean of $|B \cap \mathbb{X}|$ (the square root of the denominator) than $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$ and $\widehat{\mathrm{L}}_{\kappa M}$. Estimates for the variance in box mass and mean of box mass are not easy to extract for $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$ and are not shown in Figure 7.

Square Observation Window 40 **Units Wide** All the estimators applied to binary maps with the small square observation window showed very large bias. No estimator was able to give GBL estimates for boxes wider than 40 units as $\widehat{\mathrm{L}}_{\mathrm{GB}}$ could not place boxes of this size within the observation window and the covariance-based estimators required estimates of covariance that were not possible with such a small observation window.

The balanced pair-correlation based estimators, $\widehat{\mathrm{L}}_{gM}$ and $\widehat{\mathrm{L}}_{gI}$, applied to binary maps with the small square observation window both have a sharp rise in bias and variance at a box width of 35 units. In the case of $\widehat{\mathrm{L}}_{gI}$ this was caused by two of the 1000 realisations of $\mathbb{X}$; due to the modified denominator in $\widehat{\mathrm{L}}_{gI}$ these two realisations leaped to values on the order of 10^6 at box widths of 35. The increment size of 1 for box widths could have caused estimates from both realisations to appear to sharply rise at the same box width. It is a bit puzzling that the bias and variance of $\widehat{\mathrm{L}}_{gM}$ experienced a sharp rise at exactly the same box width as $\widehat{\mathrm{L}}_{gI}$.

Except for the sharp rise in the variance of $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$, the unbalanced covariance estimator, $\widehat{\mathrm{L}}_C$, had the largest variance for boxes wider than 12, whilst $\widehat{\mathrm{L}}_{\mathrm{GB}}$ had the smallest variance and largest bias for boxes wider than 12. A box and whisker plot of the distribution of $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_C$, $\widehat{\mathrm{L}}_{\kappa H}$ and $\widehat{\mathrm{L}}_{gI}$ is shown in Figure 8. Box widths wider than 12 appear to correspond with $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_C$ producing estimates that were below 1 (occassional estimates from $\widehat{\mathrm{L}}_{\kappa H}$ also apppear below 1 for boxes wider than 28).

Note that the apparent divergence between estimators at box width of about 12 in Figure 5 is due to the plotted scale of the GBL estimates; the pointwise variance of GBL estimators relative to the pointwise variance of $\widehat{\mathrm{L}}_{\kappa H}$ have large differences for boxes with widths smaller than 12 (Figure 6 lower right).

Partial Occlusions of the Study Region The bias and variance of the balanced covariance-based estimators were not noticeably affected by the occlusions that covered 50% or less of the study region with pointwise variance increasing by a factor of 0.5 at most. The pattern of occlusions covering 70% of Z increased the variance of these balanced covariance-based estimators by a factor of 1.6 for some boxes. The occlusion pattern covering 90% of Z caused a much larger increase in the bias and variance of these estimators, with variance increased by a factor of 5.7 for some boxes.

Relative to the balanced covariance-based estimators, $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was greatly affected by any level of occlusions, for example the pattern of occlusion covering 50% of Z increased the variance of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ by a factor of 14 for some box widths and decreased the box widths for which $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was defined to between 0 and less than 22.

In most cases as box size increased the mean of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ dropped to one and the variance of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ first increased above the variance of the other estimators, and then eventually dropped to zero. This occured, for example, for the study region with 31% occlusion for box widths starting from 10 with mean $\widehat{\mathrm{L}}_{\mathrm{GB}}$ estimates dropping to one for a box width of 27. This behaviour is explained by the reduced box locations available to $\widehat{\mathrm{L}}_{\mathrm{GB}}$ for larger box sizes in complicated observation windows: As box width increases the set of box centre locations used by $\widehat{\mathrm{L}}_{\mathrm{GB}}$, which is $W \ominus \check{B}$ where B is the box (see (15) and (16)), becomes comprised of smaller (and eventually fewer) regions. Boxes located with centres very close to each other overlap almost entirely, and it follows that the second moment box mass estimates for $\widehat{\mathrm{L}}_{\mathrm{GB}}$ (16) behave like estimates from increasingly correlated samples. When the box is sufficiently large that $W \ominus \check{B}$ is comprised by a single small region then the observed variance in box mass within any binary map is close to zero as the mass of boxes centred in $W \ominus \check{B}$ is highly correlated, and the resulting $\widehat{\mathrm{L}}_{\mathrm{GB}}$ estimate is nearly one.

The unbalanced covariance-based estimator, $\widehat{\mathrm{L}}_{C}$, performed better than $\widehat{\mathrm{L}}_{\mathrm{GB}}$ for the partially occluded observations of Z, but had substantially higher variance than the balanced covariance-based estimators.

Further Notes Excluding estimates from the small square observation window, the variance of each of the estimators had a local minima near box widths of 15; we are not sure of the cause for this local minima. Local minima in the variance of centred covariance and pair-correlation estimators were also observed by Picka (1997).

In Figure 6 there is a jump in pointwise bias at box widths of about 12, this seems to be caused by our numerical quadrature as it corresponds to a jump in the GBL of $\mathbb{X}$, which was computed from the model parameters and altering the resolution of the computed covariance of $\mathbb{X}$ created additional jumps.

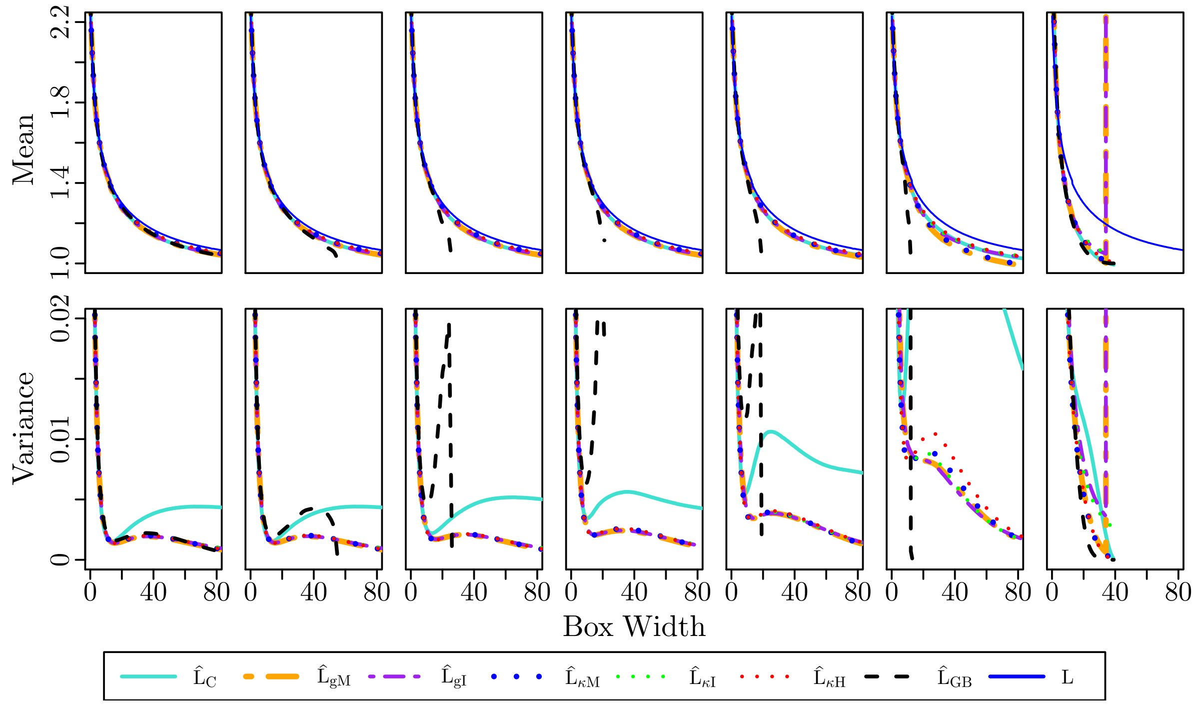

Figure 5: Results of Scenario 1. *Top:* The pointwise mean of each estimator for each observation window. *Bottom:* The pointwise variance of estimates from each estimator for each observation window. *From left:* Study region Z without occlusions; Z excluding the patterns of occlusion that covered 2%, 31%, 50%, 70% and 90% of the study region; the square observation window 40 units wide. Note that the variance and mean of the balanced covariance-based estimates were very similar in most cases and are difficult to distinguish from each other in these figures.

D.2.2 Scenario 2

The pointwise mean and variance of the GBL estimators for Scenario 2 are shown in Figure 9 with examples of the simulated binary maps. Also shown for each box width is the proportion of realisations for which there were no locations, y, such that a box

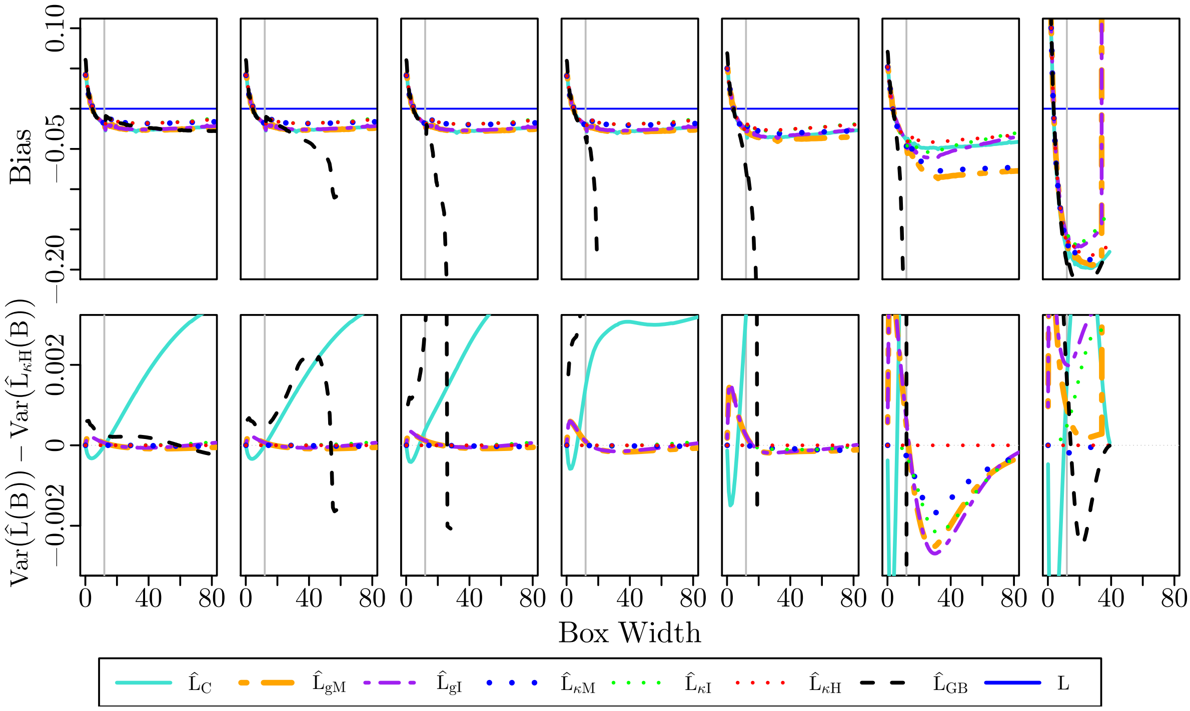

Figure 6: Pointwise bias (top) and pointwise variance relative to the pointwise variance of $\widehat{\mathrm{L}}_{\kappa H}$ (bottom). The box width of 12 is marked by a vertical grey line.

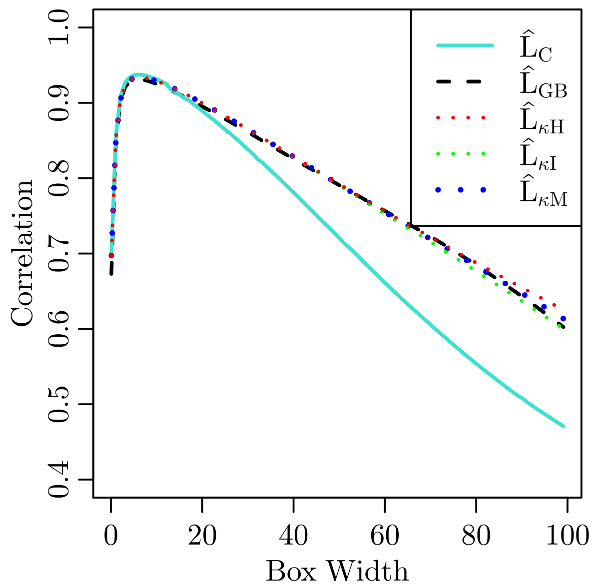

Figure 7: The Pearson correlation between estimated variance of box mass and estimated mean box mass from $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{C}$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$, and $\widehat{\mathrm{L}}_{\kappa M}$ when the study region was fully observed.

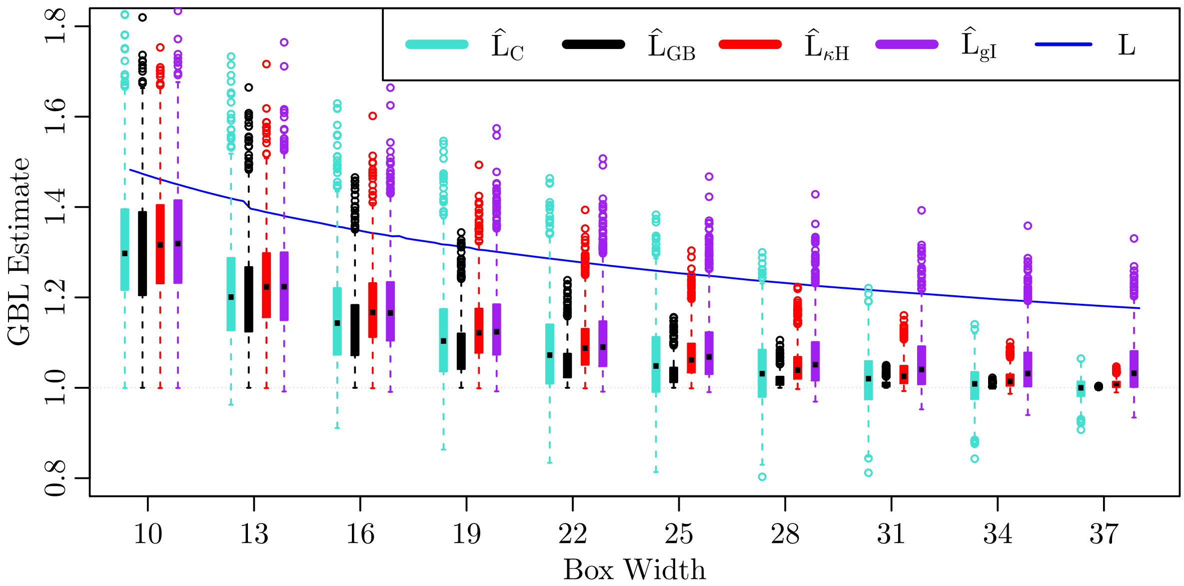

Figure 8: Distributions of estimators for the small square observation window in Scenario 1. Shown are box and whisker plots for the distributions of $\widehat{\mathrm{L}}_C$, $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{\kappa H}$ and $\widehat{\mathrm{L}}_{gI}$ at box widths of 10, 13, 16, 19, 22, 25, 28, 31, 34, and 37. The whiskers either mark the 1st and 3rd quartiles of the distributions, or are a maximum length of 1.5 of the interquartile range. The GBL of $\mathbb{X}$ is the solid blue line.

centred on y was fully contained in W (i.e. $W \ominus \check{B} = \varnothing$); for these realisations $\widehat{\mathrm{L}}_{\mathrm{GB}}$ for the given box width was undefined. The variance of the balanced covariance-based estimators observed here is 25 times smaller than the variance observed in Scenario 1.

For boxes wider than about 20 the balanced covariance-based estimators were closer to the result that $\widehat{\mathrm{L}}_{\mathrm{GB}}$ would have obtained if Z was fully observed (Figure 9, top right) and the balanced covariance-based estimators had much smaller variance than $\widehat{\mathrm{L}}_{\mathrm{GB}}$ even for small boxes. For box widths above 40 $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was increasingly unable to produce an estimate (Figure 9 centre row, right column) and was more likely to give estimates close to one, which is consistent with the low variance observed for $\widehat{\mathrm{L}}_{\mathrm{GB}}$ given large boxes in Scenario 1.

The pointwise variance $\widehat{\mathrm{L}}_C$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$, $\widehat{\mathrm{L}}_{\kappa M}$ and $\widehat{\mathrm{L}}_{\mathrm{GB}}$ have a local miminma for box widths between 1 and 12, most prominently for $\widehat{\mathrm{L}}_C$ and $\widehat{\mathrm{L}}_{\mathrm{GB}}$. We are not sure of the cause for these local minima.

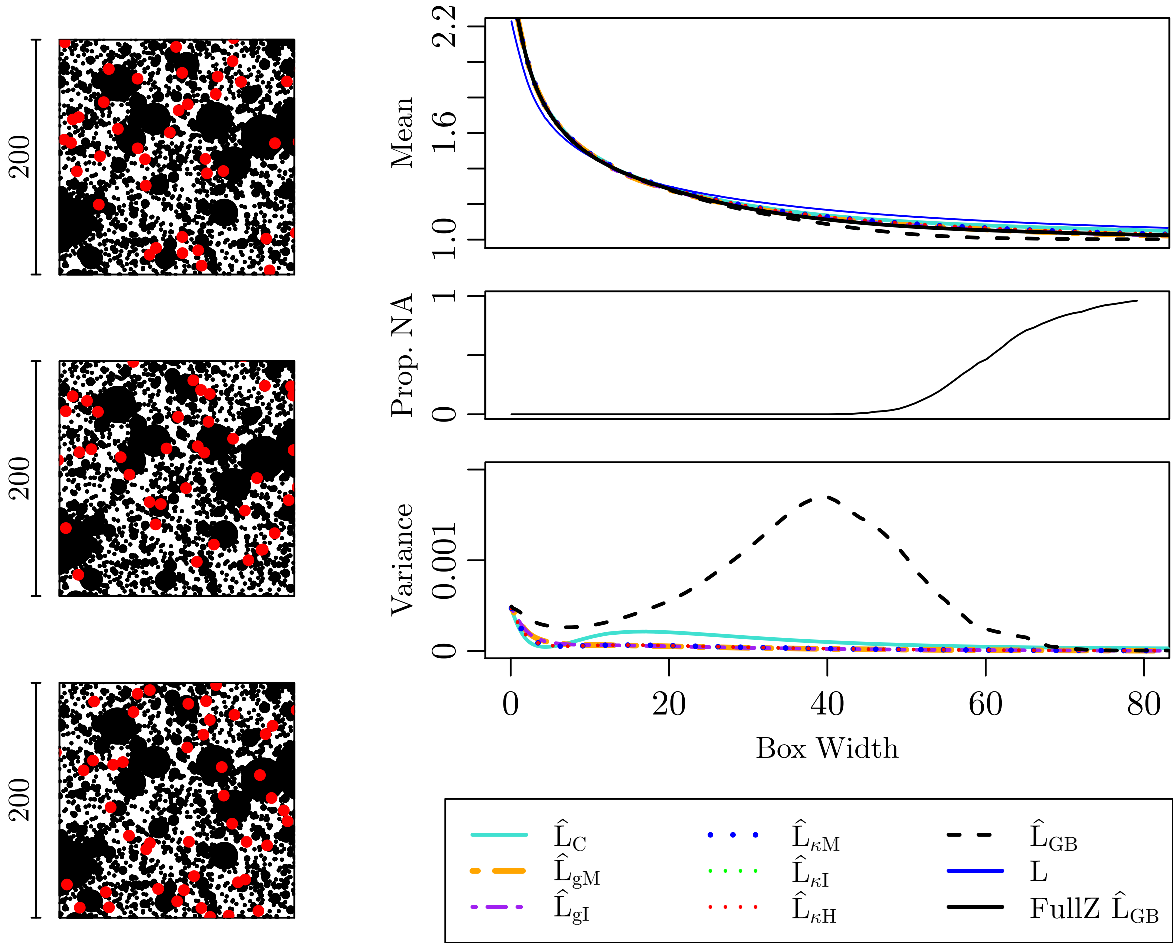

Figure 9: Results for *Scenario 2*. *Left:* The realisation, X, of $\mathbb{X}$ with example occlusion patterns. *Black:* Foreground. *White:* Background. *Red:* Occlusions. *Right: Top:* Pointwise mean of estimators. *Middle:* For each box width the proportion of realisations for which $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was undefined. *Bottom:* Pointwise variance of estimators. *In top right - solid black line:* The $\widehat{\mathrm{L}}_{\mathrm{GB}}$ estimate from X fully observed in Z.

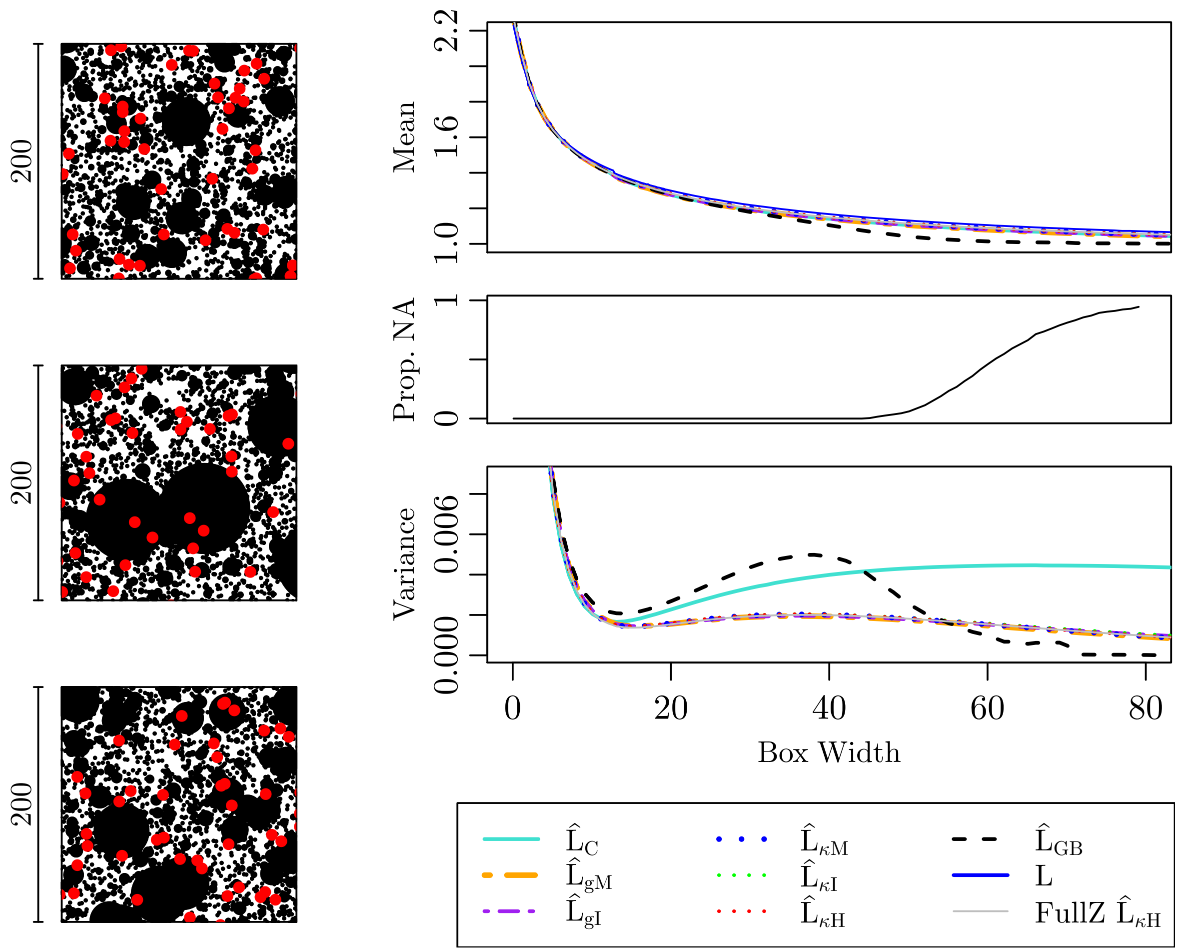

Figure 10: Results of *Scenario 3*. *Left:* Example binary maps. *Black:* Foreground. *White:* Background. *Red:* Occlusions. *Right:* *Top:* Pointwise mean of estimators. *Middle:* For each box width the proportion of realisations for which $\hat{L}_{GB}$ was undefined. *Bottom:* Pointwise variance of estimators. *Grey line:* The pointwise mean and pointwise variance of $\hat{L}_{\kappa H}$ when Z was fully observed.

D.2.3 Scenario 3

Example binary maps, and the pointwise mean and variance of GBL estimators for Scenario 3 are shown in Figure 10. Also shown for each box width is the proportion of realisations for which there were no locations, y, such that a box centred on y was fully contained in W (i.e. $W \ominus \breve{B} = \varnothing$); for these realisations $\widehat{\mathrm{L}}_{\mathrm{GB}}$ for the given box width was undefined. As with the previous studies the bias and variance of the covariance-based estimators were very similar to each other. The bias of $\widehat{\mathrm{L}}_{\mathrm{GB}}$, which diverges from the covariance-based estimators at box widths of about 27, and the variance of $\widehat{\mathrm{L}}_{\mathrm{GB}}$, which is first high and then drops to zero, are consistent with the behaviour of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ in Scenario 1. For boxes wider than approximately 50 $\widehat{\mathrm{L}}_{\mathrm{GB}}$ was often unable to produce an estimate (Figure 5 centre row, right column), which has caused some sharp features in the pointwise variance at box widths of about 70.

E Supplement to GBL Estimators Applied to Meso-Scale Forest Maps

This section describes the satellite data used Section 6.1, and then describes the method used to obtain forest presence-absence maps (Sections E.2 - E.3) and estimate GBL (Section E.4). The R code used in each step is included. Section E.5 provides R code for recreating the figures found in the main body of the text. Intermediate data is included so that the reader may skip to any section without performing earlier processing (see comments in R code).

Related files are listed in Table 2.

To execute the R code provided in this section you will need to

1. Install R: Follow the instructions on `https://www.r-project.org/` to install R.

2. Install the R packages `spatstat`, `raster`, `RcppRoll` and `maptools` from CRAN: In R run

```
> install.packages(c("spatstat", "raster", "maptools", "RcppRoll"))
```

`stationaryracsinference_0.4-01.tar.gz`	An R package containing the GBL estimators.
`satelliteimages`	A directory containing the satellite images in ERMapper raster format
`finalcloudandshadowmasks.RData`	The cloud and shadow masks used to correct the forest masks
`finalmaskedforests.RData`	The final forest masks in a format provided by R's `raster` package
`manualforrepeatinganalysis.Rnw`	The source code for this section. To recreate this document run `R Sweave` on `manualforrepeatinganalysis.Rnw`. It takes less than 10 minutes on a 3.6Ghz Ubuntu desktop.
`gbltrads.RData`	The gliding box estimates in `spatstat`'s `fv` format
`gblcs.RData`	The unbalanced covariance-based estimates in `spatstat`'s `fv` format
`gblgs.RData`	The pair-correlation based estimates in `spatstat`'s `fv` format
`gblccs.RData`	The centred covariance-based estimates in `spatstat`'s `fv` format

Table 2: Auxiliary files to this document

3. Install `stationaryracsinference`: From inside R run:

```
> install.packages("<PATH TO PACKAGE>", repos = NULL, type = "source")
```

where `<PATH TO PACKAGE>` is the path to the file `stationaryracsinference_0.4-01.tar.gz`.

E.1 Data

The satellite imagery used in Section 6.1 was part of the ARG25 (Australian Reflectance Grid 25m resolution) dataset which is a time series of calibrated multispectral 25m resolution imagery of Australia derived by Geoscience Australia from USGS's Landsat imagery (Geoscience Australia, 2015). More information on ARG25 can be found at: http://www.ga.gov.au/metadata-gateway/metadata/record/75062/ The ARG25 data is under the Creative Commons Attribution 4.0 International Licence https://creativecommons.org/licenses/by/4.0/.

Images captured from December 2015 to the end of March 2016 of a region (Geoscience

Australia tile SI50) near Albany, Western Australia, were extracted from ARG25. The seven available ARG25 images were captured on

$$\begin{gathered}\text{December 12th 2015}\\ \text{December 16th 2015}\\ \text{December 24th 2015}\\ \text{January 9th 2016}\\ \text{February 10th 2016}\\ \text{February 26th 2016}\\ \text{March 29th 2016.}\end{gathered}$$

Only the December 16th image was captured by Landsat 7; all other images were captured by Landsat 8.

A subregion that contained a multiscale forest pattern and for which the imagery contained a range of cloud occlusions was chosen to demonstrate the estimators of GBL. The extent of this subregion was (in GDA94 coordinates)

$$\begin{array}{rll}\text{Top Left:} & 504835.17\text{E} & 6195582.10\text{N}\\ \text{Bottom Right:} & 523585.17\text{E} & 6176832.10\text{N}.\end{array}$$

The imagery can be found in ERMapper format in the `satelliteimages` directory. Each image contains the following six bands: blue, green, red, near-infrared, shortwave infrared 1 and shortwave infrared 2. This corresponds to the Landsat 7 band numbers of 1, 2, 3, 4, 5 and 7, and the Landsat 8 band numbers of 2, 3, 4, 5, 6 and 7 (U. S. Geological Survey, 2018).

The following R code will load and plot the imagery in false colour.

```
> library(raster)
> #read in the ers files using the raster package
> i1208 <- brick("satelliteimages/l8region02_20151208.ers")
> i1216 <- brick("satelliteimages/l7region02_20151216.ers")
> i1224 <- brick("satelliteimages/l8region02_20151224.ers")
> i0109 <- brick("satelliteimages/l8region02_20160109.ers")
> i0210 <- brick("satelliteimages/l8region02_20160210.ers")
```

```
> i0226 <- brick("satelliteimages/l8region02_20160226.ers")
> i0329 <- brick("satelliteimages/l8region02_20160329.ers")
> #for convenience make into a list
> #chronologically ordered
> raslist.sptrl <- list(
+   s1208 = i1208,
+   s1216 = i1216,
+   s1224 = i1224,
+   s0109 = i0109,
+   s0210 = i0210,
+   s0226 = i0226,
+   s0329 = i0329)
> #plot the spectral data
> par(mfrow = c(1, 7))
> a <- lapply(raslist.sptrl, plotRGB,
+             b = 2, g = 3, r = 4, stretch = "lin")
```

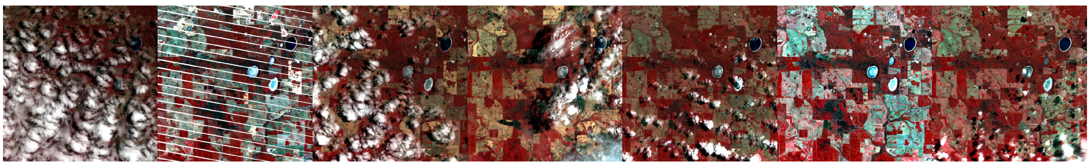

Figure 11: The satellite photographs in false colour in chronological order from left to right.

E.2 Extracting of Forest Masks

Forest masks were created from the satellite imagery by applying thresholds to combinations of spectral values. For the Landsat 8 images any pixel satisfying

$$2700 < i_4 + 3i_5 < 10000 \tag{E.1}$$

$$i_4 - i_2 > 1273 \tag{E.2}$$

was classified as forest where i_2, i_4 and i_5 were the spectral values of the green, near-infrared and shortwave infrared 1 bands respectively.

For the Landsat 7 image (December 16th's) any pixel satisfying

$$2700 < i_4 + 2i_5 < 10000 \tag{E.3}$$

$$i_4 - 2i_2 > 690 \tag{E.4}$$

was classified as forest, where i_2, i_4 and i_5 were the spectral values of the green, near-infrared and shortwave infrared 1 bands respectively.

These two sets of conditions were chosen to closely approximate the true forest cover and give similar forest presence-absence maps from each image. The following code applies the above conditions to the images.

```
> #function for classifying an individual pixel as forest in a
> #landsat 8 photograph
> l8fmasker.ppixel <- function(x) {
+     #x is a pixel, input is in b2,b4,b5
+     if (any(is.na(x))) {return(NA)}
+     else if ( (x %*% c(0, 1, 3)) > 2700 &&
+          (x %*% c(0, 1, 3)) < 10000 &&
+          x %*% c(-1, 1, 0) > 1273 ) {
+          return(TRUE)
+          }
+     else {return(FALSE)}
+ }
> #function for extracting the mask from an array of pixels
> l8fmasker <- function(x) {
+      #x is a raster object
+    l8.fm <- calc(subset(x, c(2, 4, 5)), l8fmasker.ppixel)
+    return(l8.fm)
+ }
```

```
> #function for classifying an individual pixel as forest in
> #Landsat 7 photograph
> l7fmasker.ppixel <- function(x) {
+       #x is a pixel, input is in  b2,b4,b5
+     if (any(is.na(x))) {return(NA)}
+     else if ( (x %*% c(0, 1, 2)) > 2700 &&
+          (x %*% c(0, 1, 2)) < 8000 &&
+          x %*% c(-2, 1, 0) > 690 ) {
+          return(TRUE)
+          }
+     else {return(FALSE)}
+ }
> #function for extracting the mask from an array of pixels
> l7fmasker <- function(x) {
+      #x is a raster object
+    l7.fm <- calc(subset(x, c(2, 4, 5)), l7fmasker.ppixel)
+    return(l7.fm)
+ }
> #apply the above masking functions to each image
> #Landsat 8 images
> raslist.fm.raw <- lapply(raslist.sptrl[
+                      names(raslist.sptrl) != "s1216"],
+                    l8fmasker)
> #Landsat 7 image
> raslist.fm.raw <- c(raslist.fm.raw,
+                     s1216 = l7fmasker(raslist.sptrl$s1216))
> #chronologically order list again
> raslist.fm.raw <- raslist.fm.raw[names(raslist.sptrl)]
> #plot
> par(mfrow = c(1, 7), mar = c(0, 0, 0, 0), oma = c(0, 0, 0, 0))
```

```
> a <- lapply(raslist.fm.raw, plot,
+              axes = FALSE, legend = FALSE, box = FALSE)
```

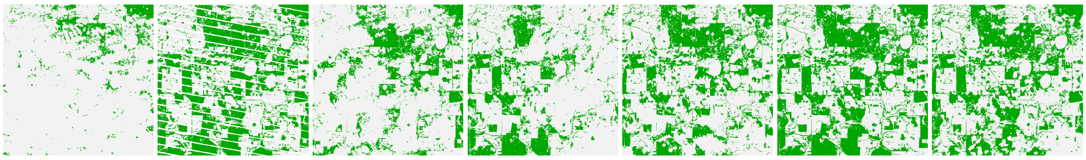

Figure 12: Raw forest masks derived from the photographs in chronological order from left to right. Green: Forest. Grey: Not-forest.

E.3 Correcting Forest Masks

The above forest masks were corrected using cloud and shadow masks, which can be can be found in the objects `cmks.final` and `sms.final` respectively in the R data file `finalcloudandshadowmasks.RData`. The cloud masks were created by thresholding of spectral values and dilating the results mask by 175m. Additional locations wispy cloud were added manually. Clouds were not noticed in the images captured on December 16th and February 26th. A mask of shadows due to clouds was created using thresholding of spectral values and dilating by $75m$. Dark lakes and other errors in the shadow mask were manually excluded before the dilation. The remaining differences between the final forest masks were relatively small compared to the size of the study region and were not corrected.

The following code loads and applies the cloud and shadow masks to the forest masks and then plots the final forest masks. Each final forest mask is a `RasterLayer` object with forest represented as a `1`, not forest represented as `0` and unobserved locations represented as `NA`. For convenience these final forest masks, `raslist.fm.m`, have been prepared earlier in `finalmaskedforests.RData`.

```
> load(file = "finalcloudandshadowmasks.RData")
> raslist.fm.m <- raslist.fm.raw
> #applying the cloud mask:
> raslist.fm.m[names(cmks.final)] <-  mapply(mask,
+                          x = raslist.fm.m[names(cmks.final)],
```

```
+                    mask = cmks.final,
+                    maskvalue = 1,
+                    updatevalue = NA,
+                    SIMPLIFY = FALSE)
> #applying the shadow mask:
> raslist.fm.m[names(sms.final)] <-  mapply(mask,
+                    x = raslist.fm.m[names(sms.final)],
+                    mask = sms.final,
+                    maskvalue = 1,
+                    updatevalue = NA,
+                    SIMPLIFY = FALSE)
> #plot masks
> par(mfrow = c(1, 7), mar = c(0, 0, 0, 0), oma = c(0, 0, 0, 0))
> a <- lapply(raslist.fm.m, plot, colNA = "gray", axes = FALSE,
+              legend = FALSE, box = FALSE)
> #save
> save(raslist.fm.m, file = "finalmaskedforests.RData")
```

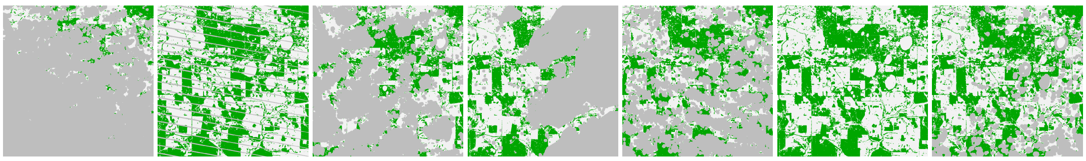

Figure 13: Final forest masks derived from the photographs in chronological order from left to right. Green: Forest. White: Not-forest. Grey: Missing data (due to cloud, cloud shadow or SLC-off).

E.4 Estimating GBL

Functions in the attached R package, `stationaryracsinference`, were used to compute GBL estimates using all covariance-based estimators and the traditional gliding box estimator. The following R code coverts the final forest masks into an appropriate format and applies the GBL estimators. We will use the `maptools` package to convert the `Raster-`

Layer objects (raslist.fm.m) into im objects (the im format comes from the spatstat package, which stationaryracsinference heavily depends on).

```
> #uncomment the following if want to skip earlier processing:
> #load("finalmaskedforests.RData")
> #
> library(spatstat)
> library(stationaryracsinference)
> library(maptools)
> #
>
> #first convert the forest masks to spatstat images
> fm.im <- lapply(raslist.fm.m, as.im)
> #
> #specify the box widths of interest
> #(implies boxes will be squares)
> sidel <- seq(25, 6000, by = 100)
> #
> #estimate GBL with the traditional estimator
> gbltrads <- mapply(gbltrad,
+                    xiim = fm.im,
+                    SIMPLIFY = FALSE,
+                    MoreArgs = list(boxwidths = sidel))
> save(gbltrads, file = "gbltrads.RData")
> #
> #estimate GBL using the unbalanced covariance-based estimator
> gblcs <- mapply(gblc,
+                   boxes = list(sidel),
+                   xiim = fm.im,
+                   SIMPLIFY = FALSE)
> names(gblcs) <- names(fm.im)
```

```
> save(gblcs, file = "gblcs.RData")
> #
> #estimate GBL using the centred covariance estimator
> #(estimating centred covariance first as this will be
> #  faster for multiple estimators)
> phats <- lapply(fm.im, coverageprob)
> ccvcsims <- lapply(fm.im, cencovariance, estimators = "all")
> allgblccsscene <- function(scenename){
+   out <- mapply(gblcc,
+         boxes = list(sidel),
+         cencovar = ccvcsims[[scenename]],
+         p = phats[[scenename]],
+         SIMPLIFY = FALSE)
+   names(out) <- names(ccvcsims[[scenename]])
+   return(out)
+ }
> gblccs <- lapply(names(phats), allgblccsscene)
> names(gblccs) <- names(phats)
> rm(ccvcsims)
> save(gblccs, file = "gblccs.RData")
> # estimates using pair-correlation
> pclnhats <- lapply(fm.im, paircorr, estimators = "all")
> allgblgsscene <- function(scenename){
+   out <- mapply(gblg,
+         boxes = list(sidel),
+         paircor = pclnhats[[scenename]],
+         SIMPLIFY = FALSE)
+   names(out) <- names(pclnhats[[scenename]])
+   return(out)
+ }
```

```
> gblgs <- lapply(names(pclnhats), allgblgsscene)
> names(gblgs) <- names(pclnhats)
> rm(pclnhats)
> save(gblgs, file = "gblgs.RData")
```

E.5 Preparing of Figures

The following R code recreates the figures in Section 6.1 that contain GBL estimates.

First prepare the estimates for easy plotting.

```
> ####
> #uncomment following if want to skip to here:
> # library(stationaryracsinference)
> # library(spatstat)
> # library(maptools)
> #a <- lapply(c("gblcs.RData", "gbltrads.RData",
> #              "gblccs.RData", "gblgs.RData"), load)
> ####
>
>
> #first convert estimates for each scene
> # to one large spatstat function object
> #(class `fv') for easy plotting
> #a function for doing this:
> joinests <- function(fvlist, ylab =  expression(hat(L))){
+   fvlist <- lapply(fvlist,
+        function(fvsingle) {fvsingle[, c(fvnames(fvsingle, ".x"),
+                            fvnames(fvsingle, ".y")), drop = FALSE]})
+   #next change the name of the function value to be the scene name
+   fvlist <- mapply(tweak.fv.entry, fvlist,
+                    new.labl = names(fvlist),
+                    new.tag = names(fvlist),
```

```
+                    MoreArgs = list(current.tag = "GBL"),
+                    SIMPLIFY = FALSE)
+   fvcomb <- do.call(cbind, fvlist)
+   fvcomb <- fv(fvcomb, argu = "s", valu = "s0226",
+                ylab = ylab,
+            labl = c("Sidelength", names(fvcomb)[-1]),
+            desc = NULL)
+   return(fvcomb)
+ }
> #applying above function:
> gblcs.fv <- joinests(gblcs, ylab = expression(hat(L)[C]))
> gbltrads.fv <- joinests(gbltrads, ylab = expression(hat(L)[GB]))
> gblccs.pickaH.fv <- joinests(
+   lapply(gblccs, "[[", "pickaH"),
+   ylab = "GBL by Centred Covaraince: Picka's H")
> gblccs.pickaint.fv <- joinests(
+   lapply(gblccs, "[[", "pickaint"),
+   ylab = "GBL by Centred Covaraince: Picka's Intrinsic")
> gblccs.mattfeldt.fv <- joinests(
+   lapply(gblccs, "[[", "mattfeldt"),
+   ylab = "GBL by Centred Covaraince: Mattfeldt-Stoyan Inspired")
> gblgs.pickaH.fv <- joinests(
+   lapply(gblgs, "[[", "pickaH"),
+   ylab = "GBL by Pair-Correlation: Picka's H")
> gblgs.pickaint.fv <- joinests(
+   lapply(gblgs, "[[", "pickaint"),
+   ylab = "GBL by Pair-Correlation: Picka's Intrinsic")
> gblgs.mattfeldt.fv <- joinests(
+   lapply(gblgs, "[[", "mattfeldt"),
+   ylab = "GBL by Pair-Correlation: Mattfeldt-Stoyan Inspired")
```

```
> #
> # make a list of the above estimates
> r2_gbl_byestimator <- list(
+       gbltrad = gbltrads.fv,
+       gblc = gblcs.fv,
+       gblcc.pickaH = gblccs.pickaH.fv,
+       gblcc.pickaint = gblccs.pickaint.fv,
+       gblcc.mattfeldt = gblccs.mattfeldt.fv,
+       gblg.pickaint = gblgs.pickaint.fv,
+       gblg.mattfeldt = gblgs.mattfeldt.fv
+       )
> #
> #
> chronoscenenames <- list(
+ s1208 = "Dec. 8th",
+ s1216 = "Dec. 16th",
+ s1224 = "Dec. 24th",
+ s0109 = "Jan. 9th",
+ s0210 = "Feb. 10th",
+ s0226 = "Feb. 26th",
+ s0329 = "Mar. 29th")
```

Now plot by estimator.

```
> pltfmla.str <- paste0('cbind(',
+             paste(names(chronoscenenames), collapse = ', '),
+             ')', " ~ ", "s")
> par(mfrow = c(2, 7),
+     oma = c(6, 3, 1.5, 0.5),
+     cex.axis = 1.4,
+     mar = c(0.5, 0.5, 1, 0.5),
+     xpd = FALSE)
```

```
> #function to build plot region and axis
> preptoplot <- function(){
+   plot.new()
+   plot.window(xlim = c(0, 5000), ylim = c(0 + 1, 1.3 + 1))
+   #axis(1, at = seq(0, 5000, by = 1000), labels = seq(0,5, by = 1))
+   box()
+ }
> #function for plotting lines from an fv object containing GBL estimates
> linesgblest.fv <- function(fvobj){
+   plot(add = TRUE, fvobj, fmla = pltfmla.str, lty = c(rep("solid", 3),
+     rep("dotdash", 2), rep("longdash", 2)), lwd = 2,
+     col = rainbow(7)[c(1, 2, 6, 5, 4, 3, 7)])
+ }
> # now using above functions to do the plots
> preptoplot()
> title(main = expression(hat(L)[GB]), line = 1, xpd = NA)
> linesgblest.fv(gbltrads.fv)
> axis(2, at = seq(0, 1.3, by = 0.1) + 1,
+   labels = c(seq(0, 1.3, by = 0.1)) + 1)
> mtext("GBL Estimate", side = 2,
+           outer = TRUE, adj = 0.8, line = 1.5)
> preptoplot()
> title(main = expression(hat(L)[C]), line = 1, xpd = NA)
> linesgblest.fv(gblcs.fv)
> preptoplot()
> title(main = expression(hat(L)[kappa * H]), line = 1, xpd = NA)
> linesgblest.fv(gblccs.pickaH.fv)
> preptoplot()
> title(main = expression(hat(L)[kappa * I]), line = 1, xpd = NA)
> linesgblest.fv(gblccs.pickaint.fv)
```

```
> preptoplot()
> title(main = expression(hat(L)[kappa * M]), line = 1, xpd = NA)
> linesgblest.fv(gblccs.mattfeldt.fv)
> preptoplot()
> title(main = expression(hat(L)[g * I]), line = 1, xpd = NA)
> linesgblest.fv(gblgs.pickaint.fv)
> preptoplot()
> title(main = expression(hat(L)[g * M]), line = 1, xpd = NA)
> linesgblest.fv(gblgs.mattfeldt.fv)
> #parameters for plot
> pltfmla.str <- paste0('cbind(',
+           paste(names(chronoscenenames[-6]), collapse = ', '),
+           ')', " - s0226", " ~ ", "s")
> #plot region and furniture setup
> preptoplot <- function(){
+   plot.new()
+   plot.window(xlim = c(0, 5000), ylim = c(-0.6, 0.6))
+   axis(1, at = seq(0, 5000, by = 1000), labels = seq(0, 5, by = 1))
+   box()
+   abline(h = 0, lty = "dashed")
+ }
> #For adding lines from an fv object containing GBL estimates
> linesgblest.fv <- function(fvobj){
+   plot.fv(add = TRUE,
+     fvobj,
+     fmla = pltfmla.str,
+     lty = c(rep("solid", 3), rep("dotdash", 2), rep("longdash", 1)),
+     lwd = 2,
+     col = rainbow(7)[c(1, 2, 6, 5, 4, 3, 7)][-6])
+ }
```

```
> #plot difference of covaiance-based estimates
> #to s0226 covariance-based estimate
> preptoplot()
> linesgblest.fv(gbltrads.fv)
> mtext("Box Width (km)", side = 1,
+           outer = TRUE, line = 1.5)
> mtext("Difference to Feb. 26th", side = 2,
+           outer = TRUE, adj = 0.1, line = 1.5)
> #title(main = expression(hat(L)[GB]))
> axis(2, at = seq(-0.6, 0.6, by = 0.2),
+      labels = c(seq(-0.6, 0.6, by = 0.2)))
> preptoplot()
> #title(main = expression(hat(L)[C]))
> linesgblest.fv(gblcs.fv)
> preptoplot()
> #title(main = expression(hat(L)[kappa * H]))
> linesgblest.fv(gblccs.pickaH.fv)
> preptoplot()
> #title(main = expression(hat(L)[kappa * I]))
> linesgblest.fv(gblccs.pickaint.fv)
> preptoplot()
> #title(main = expression(hat(L)[kappa * M]))
> linesgblest.fv(gblccs.mattfeldt.fv)
> preptoplot()
> #title(main = expression(hat(L)[g * I]))
> linesgblest.fv(gblgs.pickaint.fv)
> preptoplot()
> #title(main = expression(hat(L)[g * M]))
> linesgblest.fv(gblgs.mattfeldt.fv)
> par(fig = c(0, 1, 0, 1),
```

```
+       oma = c(0, 0, 0, 0),
+       mar = c(0, 0, 0, 0),
+       new = TRUE)
> plot(0, 0, type = "n", bty = "n", xaxt = "n", yaxt = "n")
> legend("bottom",
+          xpd = TRUE,
+          ncol = 4,
+          seg.len = 3,
+          legend = chronoscenenames,
+          col = rainbow(7)[c(1, 2, 6, 5, 4, 3, 7)],
+          lty = c(rep("solid", 3), rep("dotdash", 2), rep("longdash", 2)),
+          lwd = 2
+          )
```

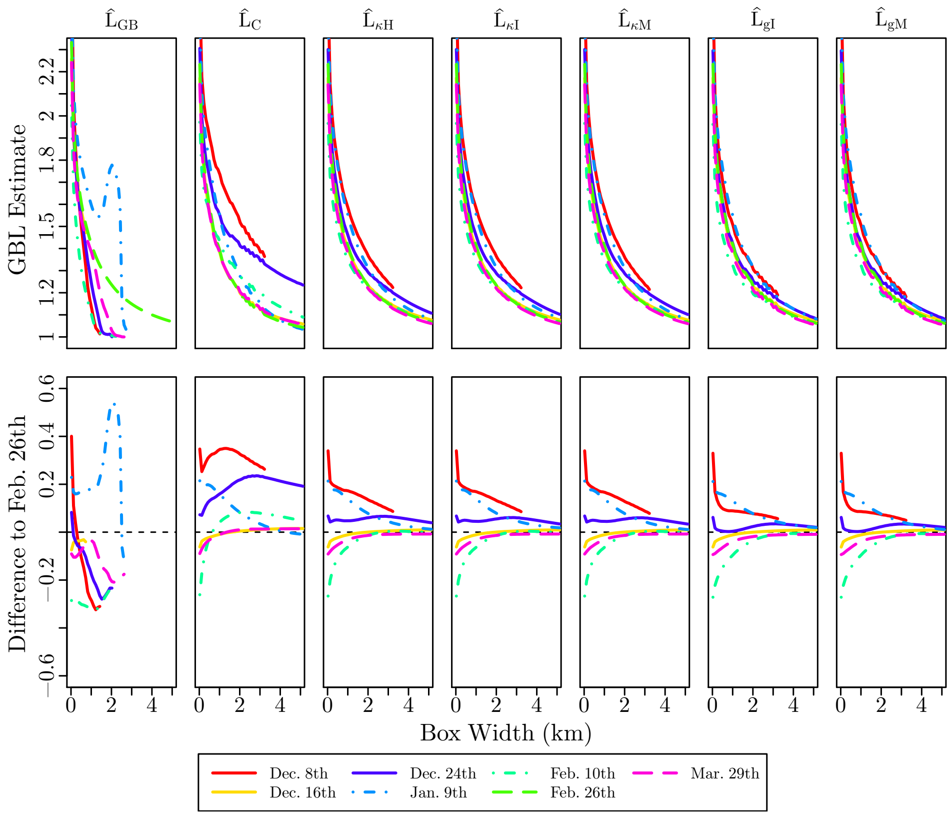

Figure 14: Results from each GBL estimator applied to the forest maps in Figure 5. *Top:* GBL estimates given square boxes. *Bottom:* The differences between estimates from the February 26 scene and estimates from the other scenes.

Now plotting estimates by scene:

```
> #parameters for plot
> par(mfrow = c(2, 7),
+     oma = c(5, 3, 1.5, 0.5),
+     mar = c(1, 0.5, 1, 0.5),
+     cex.axis = 1.4,
+     xpd = FALSE)
> #
> #set line styles
> linecols <- c(
+   gblg.none = "turquoise",
+   gblc = "turquoise",
+   gbltrad = "black",
+   gblcc.pickaH = "red",
+   gblcc.pickaint = "green",
+   gblcc.mattfeldt = "blue",
+   gblg.pickaint = "purple",
+   gblg.mattfeldt = "orange"
+ )
> linewd <- c(
+   gblg.none = 2,
+   gblc = 2,
+   gbltrad = 2,
+   gblcc.pickaH = 2,
+   gblcc.pickaint = 2,
+   gblcc.mattfeldt = 3,
+   gblg.pickaint = 2,
+   gblg.mattfeldt = 3
+ )
> linetypes <- c(
```

```
+   gblg.none = "solid",
+   gblc = "solid",
+   gbltrad = "dashed",
+   gblcc.pickaH = "dotted",
+   gblcc.pickaint = "dotted",
+   gblcc.mattfeldt = "dotted",
+   gblg.pickaint = "twodash",
+   gblg.mattfeldt = "twodash"
+ )
> linenames <- c(
+   gblg.none = expression(hat(L)[C]),
+   gblc = expression(hat(L)[C]),
+   gbltrad = expression(hat(L)[GB]),
+   gblcc.pickaH = expression(hat(L)[kappa * H]),
+   gblcc.pickaint = expression(hat(L)[kappa * I]),
+   gblcc.mattfeldt = expression(hat(L)[kappa * M]),
+   gblg.pickaint = expression(hat(L)[g * I]),
+   gblg.mattfeldt = expression(hat(L)[g * M]),
+   gbl.th = expression(L)
+ )
> plotstyles <- data.frame(col = linecols,
+                          lty = linetypes,
+                          lwd = linewd,
+                          stringsAsFactors = FALSE)
> lines.special <- function(fvobject, scenename, abbrv){
+   do.call(lines, args = c(list(
+     x = fvobject$s,
+     y = fvobject[[scenename]]),
+     plotstyles[abbrv, c("col", "lty", "lwd")])
+     )
```

```
+ }
> #a customised plot function
> plotsceneGBL <- function(scenename, scenelabel){
+   plot.new()
+   plot.window(xlim = c(0, 5000), ylim = c(0 + 1, 1.3 + 1))
+   title(main = scenelabel, font.main = 1, line = 0.5, xpd = NA)
+   axis(1, at = seq(0, 5000, by = 1000), labels = seq(0, 5, by = 1))
+   if (scenename == "s1208"){axis(2, at = seq(0, 1.3, by = 0.1) + 1,
+                           labels = c(seq(0, 1.3, by = 0.1)) + 1)
+   }
+     box()
+     out <- lapply(names(r2_gbl_byestimator),
+            function(x) lines.special(r2_gbl_byestimator[[x]],
+                                      scenename,
+                                      x))
+ }
> lines.special.loglog <- function(fvobject, scenename, abbrv){
+   do.call(lines, args = c(list(
+     x = log(fvobject$s),
+     y = log(fvobject[[scenename]])),
+     plotstyles[abbrv, c("col", "lty", "lwd")])
+     )
+ }
> #customised plot function for log log
> plotsceneGBL.loglog <- function(scenename, scenelabel){
+   plot.new()
+   plot.window(xlim = log(c(100, 5000)), ylim = log(c(1, 2.5)))
+   #title(main = scenelabel, font.main = 1)
+   axis(1, at = log(c(100, 200, 400, 800, 1600, 3200)),
+          labels = c(0.1, 0.2, 0.4, 0.8, 1.6, 3.2))
```

```
+   if (scenename == "s1208"){axis(2,
+          at = log(1.2 ^ (0:5)),
+          labels = round(1.2 ^ (0:5), 2))
+   }
+     box()
+     out <- lapply(names(r2_gbl_byestimator),
+            function(x) lines.special.loglog(r2_gbl_byestimator[[x]],
+                                          scenename,
+                                          x))
+ }
> plotsceneGBL("s1208", "Dec. 8th")
> plotsceneGBL("s1216", "Dec. 16th")
> plotsceneGBL("s1224", "Dec. 24th")
> plotsceneGBL("s0109", "Jan. 9th")
> plotsceneGBL("s0210", "Feb. 10th")
> plotsceneGBL("s0226", "Feb. 26th")
> plotsceneGBL("s0329", "Mar. 29th")
> plotsceneGBL.loglog("s1208", "")
> plotsceneGBL.loglog("s1216", "")
> plotsceneGBL.loglog("s1224", "")
> plotsceneGBL.loglog("s0109", "")
> plotsceneGBL.loglog("s0210", "")
> plotsceneGBL.loglog("s0226", "")
> plotsceneGBL.loglog("s0329", "")
>     mtext("Box Width (km)", side = 1,
+           outer = TRUE, line = 1.5)
>     mtext(c("L(B) Estimate", "L(B) Estimate"), side = 2,
+           outer = TRUE, line = 1.5,
+                   adj = c(0.2, 0.78))
> par(fig = c(0, 1, 0, 1),
```

```
+     oma = c(0, 0, 0, 0),
+     mar = c(0, 0, 0, 0),
+     new = TRUE)
> plot(0, 0, type = "n", bty = "n", xaxt = "n", yaxt = "n")
> legend("bottom",
+        xpd = TRUE,
+        ncol = 7,
+        seg.len = 3,
+        legend = linenames[names(r2_gbl_byestimator)],
+        lwd = plotstyles[names(r2_gbl_byestimator), "lwd"],
+        col = plotstyles[names(r2_gbl_byestimator), "col"],
+        lty = plotstyles[names(r2_gbl_byestimator), "lty"]
+        )
```

E.6 R Session Information

- R version 3.5.1 (2018-07-02), `x86_64-pc-linux-gnu`
- Locale: `LC_CTYPE=en_AU.UTF-8`, `LC_NUMERIC=C`, `LC_TIME=en_AU.UTF-8`, `LC_COLLATE=en_AU.UTF-8`, `LC_MONETARY=en_AU.UTF-8`, `LC_MESSAGES=en_AU.UTF-8`, `LC_PAPER=en_AU.UTF-8`, `LC_NAME=C`, `LC_ADDRESS=C`, `LC_TELEPHONE=C`, `LC_MEASUREMENT=en_AU.UTF-8`, `LC_IDENTIFICATION=C`
- Running under: `Ubuntu 16.04.5 LTS`
- Matrix products: default
- BLAS: `/usr/lib/libblas/libblas.so.3.6.0`
- LAPACK: `/usr/lib/lapack/liblapack.so.3.6.0`
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils

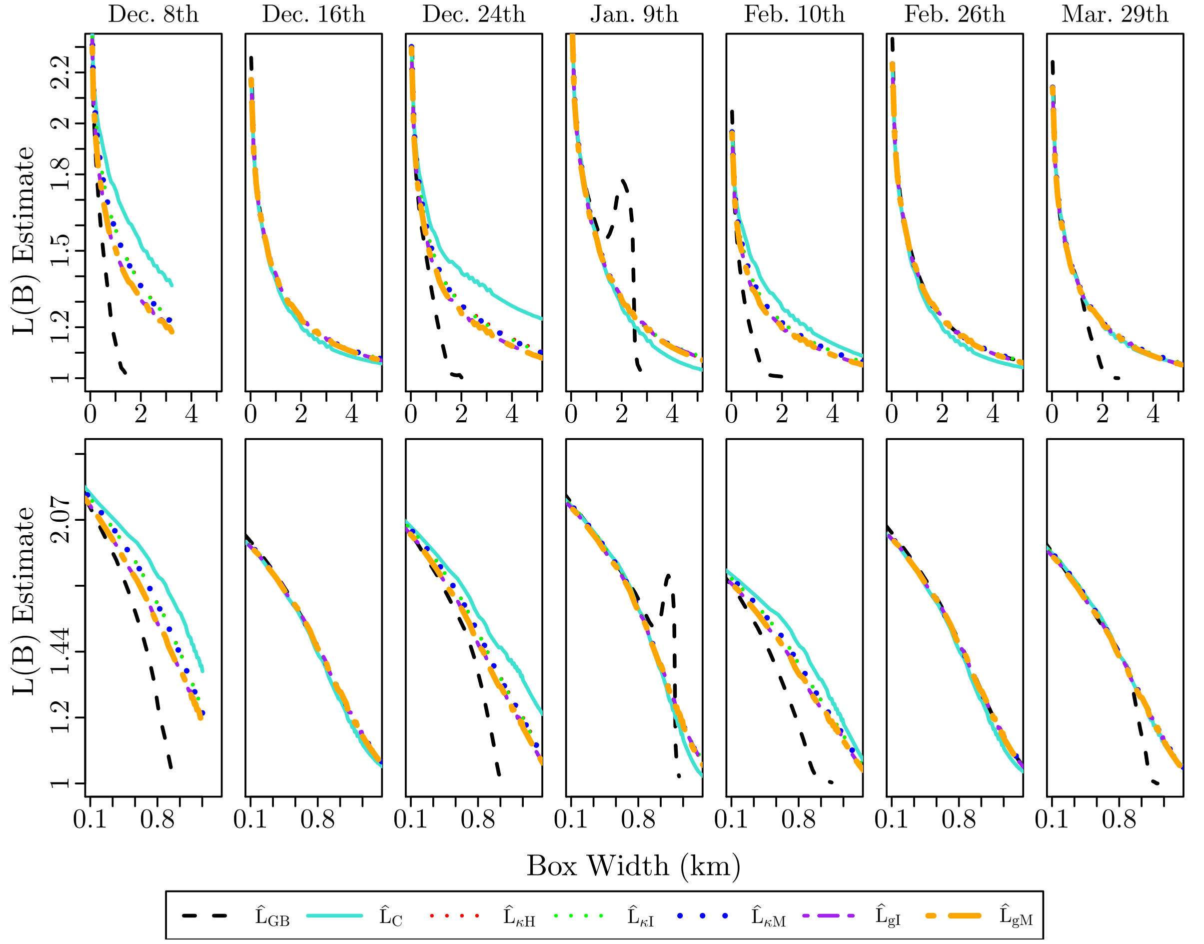

Figure 15: GBL estimates from each map in Figure 5 given square boxes. *Top:* Linear axes. *Bottom:* Log-log axes.

- Other packages: devtools 1.13.6, extrafont 0.17, maptools 0.9-3, nlme 3.1-137, raster 2.6-7, rpart 4.1-13, sp 1.3-1, spatstat 1.56-1.006, spatstat.data 1.3-1, stationaryracsinference 0.4-01, testthat 2.0.0

- Loaded via a namespace (and not attached): abind 1.4-5, compiler 3.5.1, deldir 0.1-15, digest 0.6.15, extrafontdb 1.0, foreign 0.8-71, goftest 1.1-1, grid 3.5.1, lattice 0.20-35, magrittr 1.5, Matrix 1.2-14, memoise 1.1.0, mgcv 1.8-24, polyclip 1.9-1, R6 2.2.2, Rcpp 0.12.18, RcppRoll 0.3.0, rgdal 1.3-3, rlang 0.2.1, Rttf2pt1 1.3.7, spatstat.utils 1.9-0, tensor 1.5, tools 3.5.1, withr 2.1.2

F GBL Estimators applied to Local Scale Tree Canopy Maps

An indication of the computational speed of the balanced covariance-based estimators compared to $\widehat{\mathrm{L}}_{\mathrm{GB}}$ is shown in Figure 16 for the tree canopy maps described in Section 6.2 (see Appendix G for discussion of asymptotic computation time). The estimates were computed with functions in the attached `stationaryracsinference` R package using Intel Xeon 2.20 GHz cores on a Linux machine with 65 GB of RAM. Square boxes of widths up to 1000 pixels were requested with a greater frequency at small widths and 125 different box sizes in total.

Shown in Figure 17 are GBL estimates from $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{C}$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$, $\widehat{\mathrm{L}}_{\kappa M}$, $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$ applied to the tree canopy maps described in Section 6.2. The estimates are shown transformed by $T(\hat{\mathrm{L}}) = (\hat{\mathrm{L}} - 1)\hat{p}/(1 - \hat{p})$, which standardises the estimates to have a value of 1 and 0 for, respectively, arbitrarily small and arbitrarily large boxes. This transformation may be justified by considering a stationary RACS with a covariance $C(\mathbf{v})$ that is continuous at the origin and is such that $\int_{\mathbb{R}^d} C(\mathbf{v}) - p^2 dv$ is finite. Given a box rA that is a compact set scaled by r then for small r,

$$\int_{\mathbb{R}^d} \gamma_{rA}(\mathbf{v}) C(\mathbf{v})\, dv \approx p \int_{\mathbb{R}^d} \gamma_{rA}(\mathbf{v}) dv,$$

and for large r,

$$\int_{\mathbb{R}^d} \gamma_{rA}(\mathbf{v}) C(\mathbf{v})\, dv \approx |rA| \int_{\mathbb{R}^d} C(\mathbf{v}) - p^2 dv + p^2 \int_{\mathbb{R}^d} \gamma_{rA}(\mathbf{v}) dv.$$

It follows from (22) and $\int_{\mathbb{R}^d} \gamma_{rA}(\mathbf{v}) dv = |rA|^2$ that $\lim_{r\to 0} \mathrm{L}(rA) = 1/p$ and $\lim_{r\to\infty} \mathrm{L}(rA) = 1$. If the RACS is mixing then the results of GBL estimators applied to an observation window expanding in all directions (e.g. a ball with increasingly large radius) will also converge to $1/p$ and 1 for small and large r respectively.

For some parcels the estimates given by $\widehat{\mathrm{L}}_{C}$, $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$ contain large vertical features at a box width of 25 and after standarisation a number of estimates are negative for larger box widths. This is briefly investigated below and seems related to discrepancies between these estimates and the estimates from the complement binary map.

Relation to Estimates from the Complement Map The GBL of a stationary RACS, $\mathbb{X}$, is related to the the GBL of the complement of $\mathbb{X}$, denoted $\mathrm{L}^c(B)$, by

$$\mathrm{L}(B) = \left(\frac{1-p}{p}\right)^2 \mathrm{L}^c(B) - \left(\frac{1-p}{p}\right)^2 + 1, \tag{F.1}$$

where, as usual, p is the coverage probability of $\mathbb{X}$. If we use a superscript 'c' to denote an estimator applied to the binary map after swapping foreground with background then the estimators $\widehat{\mathrm{L}}_{\mathrm{GB}}$, $\widehat{\mathrm{L}}_{\kappa H}$, $\widehat{\mathrm{L}}_{\kappa I}$ and $\widehat{\mathrm{L}}_{\kappa M}$ are such that

$$\widehat{\mathrm{L}}_{\mathrm{GB}}(B) = \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}^c_{\mathrm{GB}}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1, \tag{F.2}$$

$$\widehat{\mathrm{L}}_{\kappa H}(B) = \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}^c_{\kappa H}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1, \tag{F.3}$$

$$\widehat{\mathrm{L}}_{\kappa I}(B) = \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{\kappa I}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1, \tag{F.4}$$

$$\widehat{\mathrm{L}}_{\kappa M}(B) = \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{\kappa M}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1, \tag{F.5}$$

repectively, and the estimators $\widehat{\mathrm{L}}_C$, $\widehat{\mathrm{L}}_{gI}$ and $\widehat{\mathrm{L}}_{gM}$ are such that,

$$\widehat{\mathrm{L}}_C(B) \neq \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}^c_C(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1 \tag{F.6}$$

$$\widehat{\mathrm{L}}_{gI}(B) \neq \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{gI}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1 \tag{F.7}$$

$$\widehat{\mathrm{L}}_{gM}(B) \neq \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{gM}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1. \tag{F.8}$$

Standardised results of $\widehat{\mathrm{L}}_{gI}$ and $\left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{gI}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1$ are shown for some parcels in Figure 18. It appears that parcels that have better behaved $\widehat{\mathrm{L}}_{gI}$ results also have little discrepancy between the $\widehat{\mathrm{L}}_{gI}$ results and the $\left(\frac{1-\hat{p}}{\hat{p}}\right)^2 \widehat{\mathrm{L}}_{gI}^{c}(B) - \left(\frac{1-\hat{p}}{\hat{p}}\right)^2 + 1$ results.

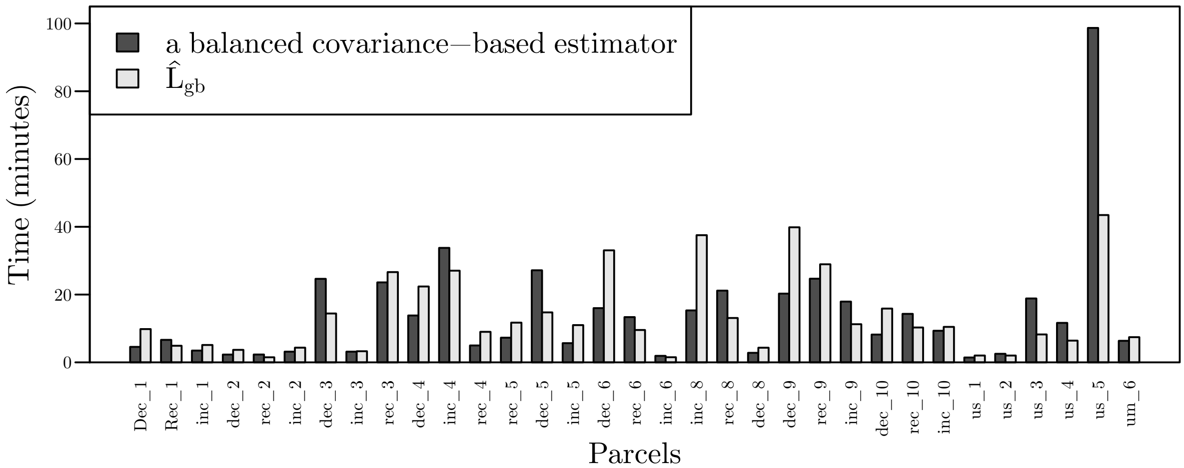

Figure 16: Computation time to estimate GBL for each parcel. *Dark:* Represents the time taken by any of the balanced covariance-based estimators. *Light:* Time taken by $\widehat{L}_{GB}$.

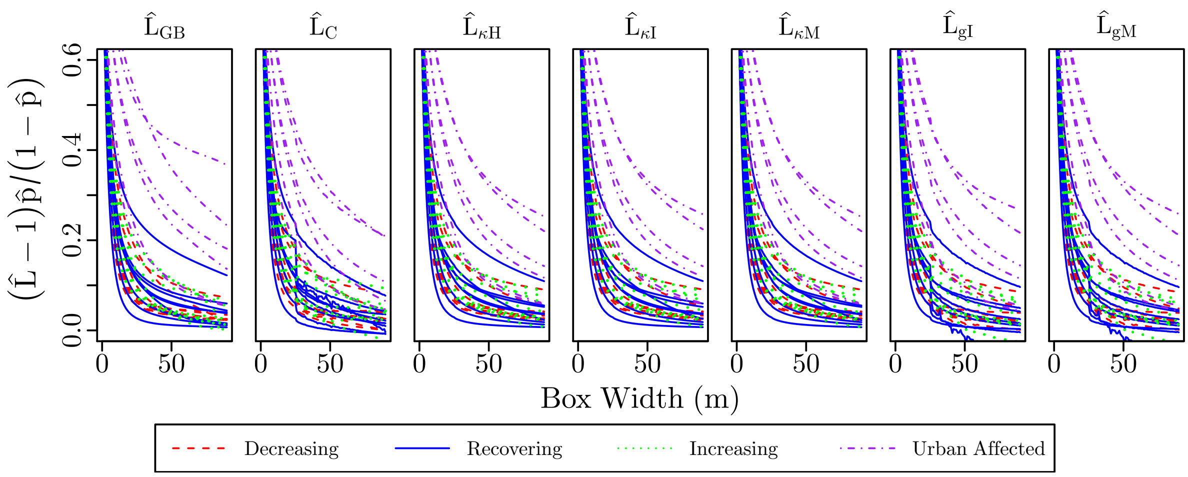

Figure 17: Estimated GBL for all 33 parcels using each available estimator. Estimates are shown standardised by the transformation $T(\hat{L}) = (\hat{L} - 1)\hat{p}/(1 - \hat{p})$.

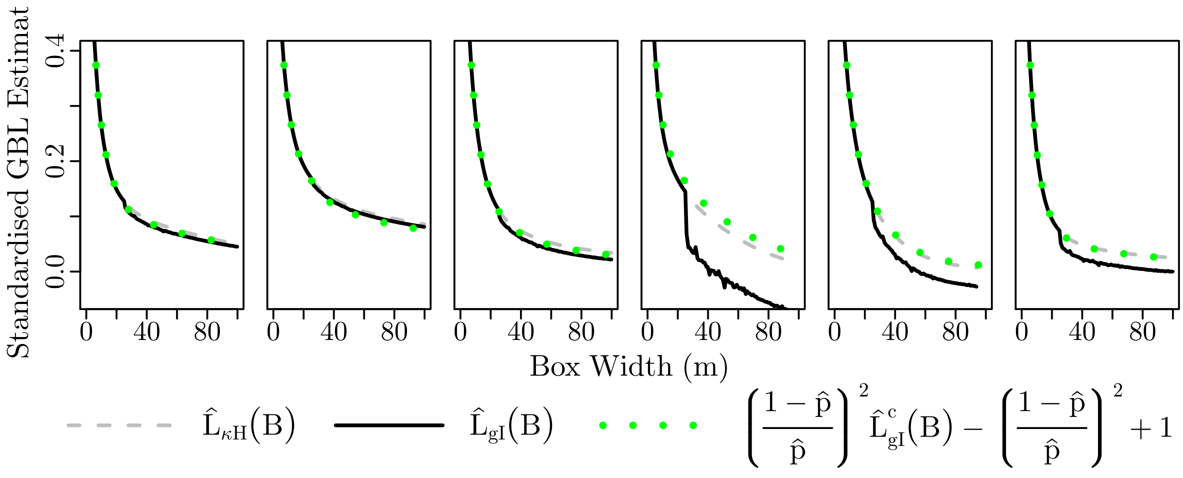

Figure 18: Results of $\widehat{\mathrm{L}}_{gI}$ and $\left(\frac{1-\hat{p}}{\hat{p}}\right)^2\widehat{\mathrm{L}}^c_{gI}(B)-\left(\frac{1-\hat{p}}{\hat{p}}\right)^2+1$ for selected parcels. Each is standardised by the transformation $T\big(\hat{\mathrm{L}}(B)\big)=(\hat{\mathrm{L}}(B)-1)\hat{p}/(1-\hat{p})$.

G Asymptotic Computational Complexity

In the attached R package, `stationaryracsinference`, the method for computing $\widehat{\mathrm{L}}_{\mathrm{GB}}$ for a given square box B used $O(n)$ box locations, where n was the number of pixels in the binary map, and required $O(\sqrt{b})$ operations, where b was the number of pixels within B, to calculate the box mass for each successive location (the number of operations is proportional to the perimeter of the box). The asymptotic complexity of $\widehat{\mathrm{L}}_{\mathrm{GB}}$ could thus be written $O(n)O(\sqrt{b})$.

To compute $\widehat{\mathrm{L}}_C(B)$ the total operations for a single box size was $O(n\log(n)) + O(b\log(b))$ as each piece of the $\widehat{\mathrm{L}}_C$ algorithm used the following number of operations:

- The covariance estimates, $\hat{C}(\mathbf{v})$, used a spatial convolution to calculate $\gamma_{X\cap W}$ and another convolution to calculate γ_W. Division of these gave $\hat{C}(\mathbf{v})$ for all locations v within $W\oplus\breve{W}$ on a grid with spatial resolution equal to the input binary map. The convolutions were calculated using three fast Fourier transforms and an image multiplication. The multiplication was an $O(n)$ operation whilst the fast Fourier transforms were each $O(n\log(n))$ operations making estimating covariance an $O(n\log(n))$ operation.

- The coverage probability estimator, $\hat{p}$, was the mean pixel value and thus required $O(n)$ operations.

- The set covariance of the box B, $\gamma_B(\mathbf{v})$, was computed using Fourier transforms, and was thus an $O(b\log(b))$ operation. For situations that require faster computations and when B is a rectangle or disc then the set covariance can be computed analytically instead (Chiu et al., 2013, §1.7.2).
- The integral of $\gamma_B(\mathbf{v})\hat{C}(\mathbf{v})$ required $O(b)$ operations.

For the balanced estimators an additional convolution for the numerator of Picka's reduced window coverage probability estimator, $\hat{p}_R(\mathbf{v})$, defined in (8), was required. A further convolution was used to compute the denominator, γ_W, of (8), which could be avoided in the future as γ_W was already computed for $\hat{C}(\mathbf{v})$. In total computation of $\hat{p}_R(\mathbf{v})$ was thus than $O(n\log(n))$ operation and estimates of GBL using the balanced covariance-based estimators required $O(n\log(n)) + O(b\log(b))$ operations.

The covariance-based estimators will scale better than $\widehat{\mathrm{L}}_{\mathrm{GB}}$ with increasing resolution if the box sizes of interest are independent of resolution (e.g. linked to a physical understanding of the observed phenomena) as the number of pixel's in each box, b, will scale with n and computing covariance-based estimates and $\widehat{\mathrm{L}}_{\mathrm{GB}}$ estimates will be, respectively, $O(n\log(n))$ and $O(n^{3/2})$ operations. Furthermore the covariance-based estimators scale better with increasing numbers of different box sizes as the $O(n\log(n))$ operations were performed once for each binary map. However if the box size scales with resolution or the observation window extent increases then computing $\widehat{\mathrm{L}}_{\mathrm{GB}}$ will be an $O(n)$ operation and will scale better than the covariance-based estimators.

References

Baddeley, A., Rubak, E., and Turner, R. (2015), *Spatial Point Patterns: Methodology and Applications with R*, London: Chapman and Hall/CRC.

Chiu, S. N., Stoyan, D., Kendall, W. S., and Mecke, J. (2013), *Stochastic Geometry and Its Applications*, 3 edn, Chichester, United Kingdom: John Wiley & Sons.

Dàvila, E., and Parés, D. (2007), "Structure of heat-induced plasma protein gels studied by fractal and lacunarity analysis," *Food Hydrocolloids*, 21(2), 147–153.

Dàvila, E., Toldrà, M., Saguer, E., Carretero, C., and Parés, D. (2007), "Characterization of plasma protein gels by means of image analysis," *LWT - Food Science and Technology*, 40(8), 1321–1329.

Geoscience Australia (2015), Australian Reflectance Grid (ARG25) Product Description,, Technical report, Commonwealth of Australia.

Hansen, M. B., Baddeley, A. J., and Gill, R. D. (1999), "First Contact Distributions for Spatial Patterns: Regularity and Estimation," *Advances in Applied Probability*, 31(1), 15–33.

Karperien, A. (2015), "FracLac for ImageJ,". Version 2015Sep090313a9330.
URL: *http://rsb.info.nih.gov/ij/plugins/fraclac/FLHelp/Introduction.htm*

Karperien, A. L. (2005), FracLac's Advanced User Manual,, Technical report, Charles Sturt University. Version 2.0f.

Mandelbrot, B. B. (1983), *Fractals and the Geometry of Nature*, New York: W. H. Freeman and Company.

Picka, J. D. (1997), Variance-reducing modifications for estimators of dependence in random sets, Ph.D., The University of Chicago, United States – Illinois.

U. S. Geological Survey (2018), "*What are the band designations for the Landsat satellites?*".
URL: *https://landsat.usgs.gov/what-are-band-designations-landsat-satellites* Accessed: 18 September 2018.