Biomass estimation in a tropical wet forest using Fourier transforms of profiles from lidar or interferometric SAR

R. N. Treuhaft,[1] F. G. Gonçalves,[1,2] J. B. Drake,[3] B. D. Chapman,[1] J. R. dos Santos,[4] L. V. Dutra,[4] P. M. L. A. Graça,[5] and G. H. Purcell[1]

Received 24 September 2010; accepted 21 October 2010; published 9 December 2010.

[1] Tropical forest biomass estimation based on the structure of the canopy is a burgeoning and crucial remote sensing capability for balancing terrestrial carbon budgets. This paper introduces a new approach to structural biomass estimation based on the Fourier transform of vertical profiles from lidar or interferometric SAR (InSAR). Airborne and field data were used from 28 tropical wet forest stands at La Selva Biological Station, Costa Rica, with average biomass of 229 Mg-ha^{-1}. RMS scatters of remote sensing biomass estimates about field measurements were 58.3 Mg-ha^{-1}, 21%, and 76.1 Mg-ha^{-1}, 26%, for lidar and InSAR, respectively. Using mean forest height, the RMS scatter was 97 Mg-ha^{-1}, ≈34% for both lidar and InSAR. The confidence that Fourier transforms are a significant improvement over height was >99% for lidar and ≈90% for InSAR. Lidar Fourier transforms determined the useful range of vertical wavelengths to be 14 m to 100 m. **Citation:** Treuhaft, R. N., F. G. Gonçalves, J. B. Drake, B. D. Chapman, J. R. dos Santos, L. V. Dutra, P. M. L. A. Graça, and G. H. Purcell (2010), Biomass estimation in a tropical wet forest using Fourier transforms of profiles from lidar or interferometric SAR, *Geophys. Res. Lett.*, *37*, L23403, doi:10.1029/2010GL045608.

1. Introduction

[2] The global measurement of aboveground forest biomass via remote sensing is essential for understanding the Earth's carbon cycle [e.g., *Houghton et al.*, 2009]. Measurements of standing biomass can be compared over different epochs to determine biomass change and infer rates of atmospheric carbon gain or loss due to land management, disturbance, and regrowth. Biomass estimation in tropical forests is particularly important because they contain approximately half of the Earth's forested biomass, and because their characterization with remote sensing is difficult due to their density and structural complexity [*Drake et al.*, 2002; *Hajnsek et al.*, 2009; *Treuhaft et al.*, 2009]. Polarimetric radar power has historically been the most common remote sensing observation from which forest biomass has been estimated, but radar power loses sensitivity to biomass increments for biomass levels between 50–150 Mg-ha^{-1}, depending on the radar wavelength and polarization [*Lu*, 2006, and references therein]. The loss in sensitivity induces a commonly cited asymptotic "saturation" in radar power as a function of biomass, reducing radar power's efficacy for estimating biomass densities ≥150 Mg-ha^{-1}, as are often found in tropical forests. However, biomass estimation based on remotely sensed forest vertical structure from lidar [e.g., *Drake et al.*, 2002] and interferometric SAR (InSAR) [e.g., *Treuhaft et al.*, 2003], evidences much less saturation beyond 200 Mg-ha^{-1} and shows improved performance relative to that based on radar power alone.

[3] Structural biomass estimation implicitly assumes that the vertical organization of the canopy of a forest stand is correlated with its biomass. It often proceeds by the regression of field-measured biomass to remotely sensed average, median, or maximum forest height. In contrast, this paper regresses field-measured biomass to the Fourier transform [*Bracewell*, 1986] of the vertical vegetation profile from either lidar or InSAR. It first describes the lidar Fourier transform and the InSAR complex coherence, showing their formal similarity and a relationship between the vertical Fourier frequency and the InSAR baseline. It then describes field structural measurements, taken at La Selva Biological Station in 2006, from which field biomass was estimated for 28 stands, with average biomass of 229 Mg-ha^{-1} (35–457 Mg-ha^{-1}). Using lidar data taken with LVIS in 2005 [*Blair et al.*, 2006], and multibaseline InSAR data at C-band taken with AirSAR in 2004 [*Treuhaft et al.*, 2009], regression of field biomass to Fourier transforms of vertical profiles at 2–3 spatial frequencies performed better than regression to mean height for both lidar and InSAR. Broad ranges of lidar Fourier frequencies are then used to bracket the range of lidar frequencies that produce the best biomass estimation accuracy. Though outside of the scope of this work, this paper is a first step toward a strategy for combining lidar Fourier transforms and InSAR complex coherences at various spatial frequencies for space-based global biomass monitoring. Such a strategy could potentially be used on a vegetation-interferometric variant of a radar-lidar mission like NASA's DESDynI. Throughout this paper, "biomass" means aboveground forest biomass per unit area.

2. Fourier Transform of Lidar and the InSAR Complex Coherence

[4] The lidar waveform, $w(z)$, is the infrared power received at nadir from altitude z in the forest due to a nadir-injected lidar pulse [*Lefsky et al.*, 2002]. The normalized

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[2]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon, USA.
[3]USDA Forest Service, Tallahassee, Florida, USA.
[4]Instituto Nacionial de Pesquisas Espaciais, São José dos Campos, SP, Brazil.
[5]Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil.

0094-8276/10/2010GL045608

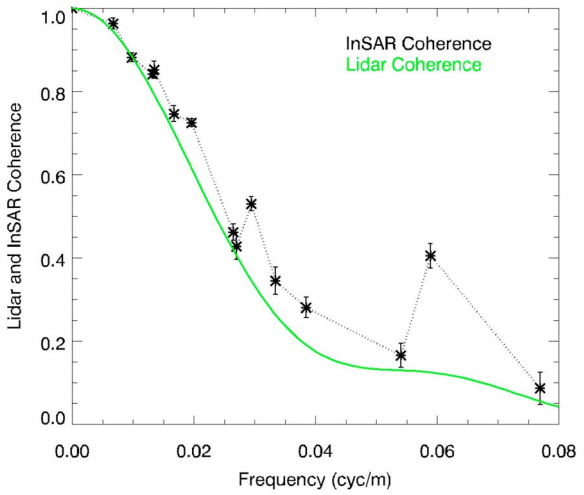

Figure 1. The lidar coherence—the amplitude of (1)—in green, as a function of vertical spatial frequency (f in (1)). The black points are InSAR coherences from 14 InSAR baselines, corresponding to vertical frequencies as in (3).

Fourier transform of $w(z)$, or lidar complex coherence $\gamma_L(f)$ at spatial frequency f, is [*Bracewell*, 1986]:

$$\gamma_L(f) = \frac{\int_0^\infty w(z)\, e^{i2\pi fz}\, dz}{\int_0^\infty w(z) dz}, \tag{1}$$

where the ground surface is assumed to be at $z = 0$, and $w(z) = 0$ at altitudes higher than the tallest trees. Qualitatively, Fourier transform amplitudes will be largest for frequencies f, when there are significant changes of forest vertical structure on spatial scales of the vertical spatial wavelength, $1/f$.

[5] For a given transmit and receive polarization combination, the contribution of vertical structure to the complex coherence for fixed-baseline InSAR—in which both ends of the baseline acquire signals simultaneously, i.e., no temporal decorrelation—can be isolated [*Rodriguez and Martin*, 1992]. For a horizontal baseline on flat terrain, the InSAR complex coherence due to vertical structure $\gamma_I(B, \theta, h, \lambda)$, depends on baseline length, B, incidence angle, θ ($\approx 35°$ for the InSAR data in this paper), radar height h, and radar wavelength, λ as [*Treuhaft et al.*, 1996]

$$\gamma_I(B, \theta, h, \lambda) = \frac{\int_0^\infty g(z, \theta, \lambda) \exp\left[\frac{i2\pi B \cos^2\theta}{\lambda h \sin\theta} z\right] dz}{\int_0^\infty g(z, \theta, \lambda)\, dz}, \tag{2}$$

where $g(z, \theta, \lambda)$ is the radar power returned from altitude z. Like $w(z)$, it depends on scatterer density and brightness, and the attenuation at z. The magnitude of either (1) or (2) will be called the lidar or InSAR "coherence," respectively. Note that the InSAR complex coherence for each baseline *is* the normalized Fourier transform of the vertical profile of return power, $g(z, \theta, \lambda)$, at the vertical frequency in (3) below.

[6] From (1) and (2), if $w(z)$ and $g(z, \theta, \lambda)$ have the same relative dependence on z, and the lidar Fourier frequency were given by

$$f = \frac{B \cos^2\theta}{\lambda h \sin\theta}, \tag{3}$$

the normalized lidar Fourier transform, $\gamma_L(f)$, would be equal to the InSAR complex coherence, $\gamma_I(B, \theta, h, \lambda)$.

[7] Figure 1 shows the lidar coherence in green, obtained by Fourier transforming the average lidar waveform of a primary forest stand, 50 m × 50 m. InSAR coherences at C-band for the same stand, corrected for instrumental effects, are shown for 14 frequencies, calculated from each of 14 baselines by an equation like (3), but accounting for the actual slope of the AirSAR baseline. Figure 7 of *Treuhaft et al.* [2009] shows InSAR profiles, corrected for a small attenuation of 0.1 dB/m extinction coefficient, and lidar and field profiles. The qualitative similarities of the lidar and InSAR coherences and estimated profiles in these figures suggest that lidar and InSAR complex coherences at C-band may have similar biomass-estimation performance. Introducing 0.3 dB/m of simulated attenuation into lidar data suggests that attenuation differences between lidar and InSAR will not affect biomass estimation by more than a few percent.

3. Field Biomass Estimation and Its Error

[8] In order to arrive at biomass estimates to be regressed to height and Fourier transforms, the mass of each tree M_{field} (kg) was estimated from measurements of the diameter at breast height or above buttress D (cm) and total height H (m), for all trees with $D > 10$ cm using the allometric equation for tropical wet forests of *Chave et al.* [2005]:

$$M_{field} = 0.0776 \times \left(\rho D^2 H\right)^{0.940}, \tag{4}$$

where ρ (g cm^{-3}) is the wood density, assumed to have a linear dependence with diameter class, with slope from Figure 2 of *Chave et al.* [2004]. Because species identification was not done, a mean value for density of 0.602 g cm^{-3}, from a broad set of measurements of wet tropical forests in Central America [*Chave et al.*, 2006], was associated with the average D measured at La Selva of 22.4 cm. The measurement of D and H is described by *Treuhaft et al.* [2009]. The biomass was calculated by adding the mass of all trees measured in a transect, and dividing by the area of the transect for field measurements, 10 m × 100 m, or 0.1 ha.

[9] The field biomass error, the error bars in Figure 2, consisted of three dominant contributions. The first involved the propagation of field errors in D, H, and ρ to biomass, using the methods of *Chave et al.* [2004]. Field measurement errors typically induced a 6% error in the stand biomass. The second contribution was that due to inaccuracies in tree mass due to the functional form of (4). This was estimated by using a correction factor as that of *Chave et al.* [2004], and was typically 9%. The third source of error arose because the field transects of 10 × 100 m were only samples of the remote sensing plots with lidar or InSAR, which were 50 × 50 m. This "counting statistical" error was estimated with a binomial statistical model to be of the order of 15–30%.

[10] From the measured biomass rates of change at La Selva [*Dubayah et al.*, 2010] and from supplementary statistics on the highest biomass rates (R. O. Dubayah, personal communication, 2010), the change in biomass due to 1–2 year differences in field and remote sensing epochs induced a ≈3% change in RMS scatters in height-based biomass regression

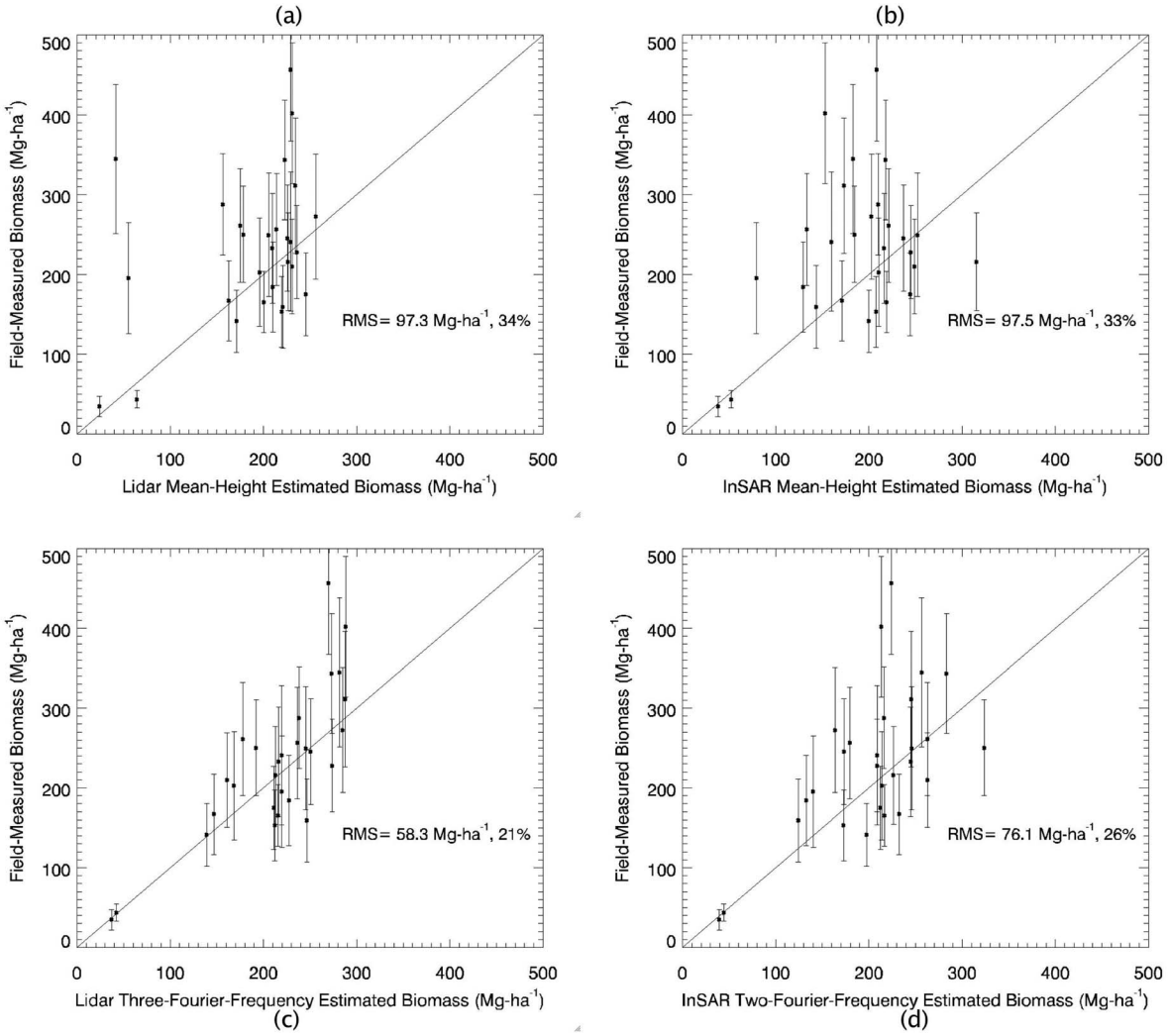

Figure 2. Field-measured biomass versus model estimates based on average height, as in (5) for (a) lidar and (b) InSAR, based on the real and imaginary parts of Fourier transforms for (c) 3 frequencies centered at 0.06 cyc/m, separated by 0.01 cyc/m for lidar and (d) 2 frequencies, 0.01 cyc/m and 0.02 cyc/m, for InSAR. The diagonal lines are y = x.

and ≈0.5% changes for Fourier regression. Forest growth was therefore not considered in the analyses that follow.

4. Biomass Estimation From Average Height

[11] For lidar and InSAR, respectively, Figures 2a and 2b show the field biomass estimates versus height-based, model biomasses given for the jth stand by

$$B_{model_j} = a + b\,\langle z\rangle_j \quad \text{with} \quad \langle z\rangle_j = \frac{\int_0^\infty z\, w_j(z)\, dz}{\int_0^\infty w_j(z) dz}, \qquad (5)$$

where the height, $\langle z\rangle_j$, is averaged over $w_j(z)$ as in (5) for Figure 2a, or its analogue averaged over $g(z, \theta, \lambda)$ for InSAR in Figure 2b. Height values used in Figures 2a and 2b are from *Treuhaft et al.* [2009]. The parameters a and b are determined for lidar and InSAR separately by minimizing χ^2 [*Hamilton*, 1964] between field and model biomasses, summed over 28 stands.

[12] The RMS scatters of the best fit model about the field biomasses are the same for lidar and InSAR at about 97 Mg-ha^{-1}, as shown in Table 1. The percent scatter is the RMS of the percentage distance from the $y = x$ line in Figure 2, and was 34% and 33%, for lidar and InSAR, respectively. The "confidence" in the first two lines of Table 1 for height-based estimation is the probability that the scatters shown result from a real correlation between field biomass and height. It was calculated as $1 - P(s)$, where $P(s)$ is the probability that the scatter, s, of column 2 (or less) could serendipitously result in the absence of any real correlation between field biomass and average height. The serendipitous probability was evaluated by randomly shuffling the field biomass measurements among the plots, thereby rendering shuffled field and height-based model biomass uncorrelated. The high confidence for both lidar and InSAR mean height entries suggests that the correlation of biomass with height is significant. The reduced χ^2 for regression to average height, which should be near 1.0, is 1.8 for both lidar and

Table 1. Performance of Biomass Estimation From Lidar and InSAR[a]

Independent Variables for Fit	RMS Scatter (Mg-ha^{-1})	RMS Percent Scatter (%)	Confidence (%)	Reduced χ^2
Mean lidar height	97.3	34	99.5	1.8
Mean InSAR height	97.5	33	99.7	1.8
3 lidar Fourier transforms	58.3	21	**99.8**	0.9
2 InSAR Fourier transforms	76.1	26	**89.8**	1.3

[a]The percent scatters in column 3 are the RMS of the percentage distances from each of the $y = x$ lines in Figure 2. The bolded values indicate conditional confidence as discussed in section 5.

InSAR, suggesting that either B_{model_j} in (5) insufficiently characterizes the biomass data, or that the biomass error bars are too small.

5. Biomass Estimation From Fourier Transforms

[13] Figures 2c and 2d show the field biomass estimates versus model biomasses which, for the *j*th stand, are based on the Fourier transform of the vertical profile:

$$B_{model_j} = a + \sum_{k=1}^{N_f} \left[a_k \times \text{real part}\left(\gamma_{\mathrm{L_j}}(\mathrm{f_k})\right) + b_k \times \text{imaginary part}\left(\gamma_{\mathrm{L_j}}(\mathrm{f_k})\right)\right] \quad (6)$$

for N_f frequencies, $\{f_k\}$. The Fourier-based model biomasses for InSAR are the same as (6), with $\gamma_{Lj}(f_k) \rightarrow \gamma_{Ij}(f_k)$, with each f_k corresponding to a baseline given by (3), but accounting for the actual slope of the AirSAR baseline and the terrain. All constant a and b parameters were chosen to minimize χ^2.

[14] The confidence for the Fourier fits in Table 1 is conceptually different from that for the height fits. In the height regression of section 4, confidence rested on the unlikelihood of realizing the observed scatters in the absence of any dynamic biomass-structure relationship. Now we want to know the probability of obtaining the reduced scatters of Figures 2c and 2d in the absence of any *further* dynamic relation regarding profiles, beyond the height relationship already established in Figures 2a and 2b. In order to assess this "conditional" confidence, we simulated arbitrarily shaped profiles for each stand, while constraining each profile to have the actual average height of that stand. Profiles were thus rendered uncorrelated with biomass, while the height correlation was preserved. The simulated data allowed determination of $P(s)$, the probability of realizing the scatter of column 2 in the Table with random profiles. The conditional confidence that Fourier fits constitute a real improvement over height fits, which is $1 - P(s)$, was >99% for groups of 3 Fourier frequencies for lidar and ≈90% for groups of 2 Fourier frequencies for InSAR. The choice of 3 and 2 frequencies for N_f for lidar and InSAR, respectively, was dictated by minimizing the RMS scatter while maximizing the conditional confidence.

[15] The optimal lidar performance was with three frequencies (vertical wavelengths), separated by 0.01 cyc/m, centered at 0.06 cyc/m (17 m). For InSAR in Figure 2d, the optimal frequencies were 0.01 and 0.02 cyc/m, with an average of 0.015 cyc/m (67 m). The InSAR frequencies were constrained to these lower values in part because the InSAR phase became unreliable for frequencies higher than 0.03 cyc/m [*Treuhaft et al.*, 2009]. The improved RMS scatters of 58.3 and 76.1 Mg/ha are shown in Table 1. The poorer InSAR performance and confidence are in part due to coherence and phase errors of 0.1 and 10°, respectively, which were estimated by adding simulated noise to lidar observations until the degraded biomass estimation performance of lidar equaled that of InSAR. This same simulation suggested that InSAR errors of 0.01 and 1° for coherence and phase could enable its performance to equal that of lidar. The reduced χ^2's were also improved for the Fourier transform model, suggesting the Fourier model better fits the field biomass.

6. Vertical Scales That Best Estimate Biomass

[16] Lidar data's broad, essentially continuous range of Fourier frequencies enabled them to determine the performance of a range of 3-frequency groups, separated by 0.01 cyc/m. Figure 3 shows the RMS estimated biomass scatter about the field measurements from lidar Fourier transforms at frequencies with the center of each group on the abscissa. The line near 90 Mg-ha^{-1} shows the RMS scatter, with 1-σ bars, arising from the random-profile simulation above, i.e., resulting from the addition of parameters and no real profile-biomass relationship beyond the height relationship. When the actual RMS biomass scatter from the data is a few bars below the line, the confidence is high. Figure 3 suggests that Fourier transforms at vertical wavelengths between ≈14 and 100 m improve upon height-only regression, but those from wavelengths outside of this range do not. Based on arguments about the near-equivalence

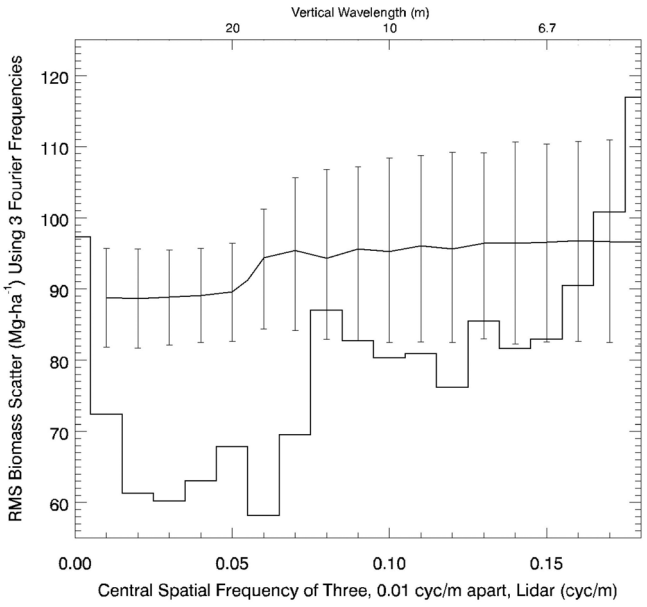

Figure 3. The RMS scatter of field biomass about model biomass using three vertical, spatial Fourier transforms of lidar profiles, with the central frequency on the x-axis. The zero-frequency entry is for the height-based fit of Figure 2a. The line near 90 Mg-ha^{-1} is the RMS scatter from simulated random profiles with the height distribution of the real data.

of γ_L and γ_I in section 2, Figure 3 might be used as a tool to select InSAR baselines via (3), if InSAR measurement errors could be reduced.

[17] It is worth noting that the average height (5) is $(2\pi i)^{-1}$ times the derivative of $\gamma_L(f)$ with respect to f in the limit as $f \rightarrow 0$. This well-known result implies that height-related metrics can be derived from very low Fourier frequencies alone. For this reason, the scatter from height in Figure 2a was placed at zero, as it can be thought of as arising from a small group of frequencies near zero. Figure 3 suggests that groups of higher frequencies, up to about 0.07 cyc/m, can potentially improve biomass estimation performance.

7. Conclusions

[18] This letter is a first attempt at applying Fourier transforms to the question: What are the most effective functions of vertical profile measurements, from lidar or InSAR, to use in tropical forest biomass estimation? It demonstrates that the performance of the frequently used low-frequency height was improved for lidar and C-band InSAR by using higher vertical frequencies, corresponding to 14–100 m vertical wavelengths. The function of the higher frequencies used, the simple sum of (6), is an empirical, first guess. Dynamical insight as to why these vertical wavelengths correlate with field biomass will, in the future, probably generate more efficient model functions than (6).

[19] This letter also suggests that, if InSAR performance could be improved, baselines corresponding to 14–100 m wavelengths might also be efficacious for biomass estimation. L-band polarimetric InSAR, with better canopy penetration and capability to isolate the ground, might perform better than the C-band InSAR reported here. It may be possible to combine lidar Fourier transforms and a few InSAR baselines in a frequency-space data fusion scheme for missions like DESDynI.

[20] **Acknowledgments.** The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

Blair, J. B., M. A. Hofton, and D. L. Rabine (2006), Processing of NASA LVIS elevation and canopy (LGE, LCE and LGW) data products, version 1.01, NASA, Washington, D. C. (Available at http://lvis.gsfc.nasa.gov)

Bracewell, R. N. (1986), *The Fourier Transform and Its Applications*, McGraw-Hill, Boston, Mass.

Chave, J., R. Condit, S. Aguilar, A. Hernandez, S. Lao, and R. Perez (2004), Error propagation and scaling for tropical forest biomass estimates, *Philos. Trans. R. Soc. B*, *359*, 409.

Chave, J., et al. (2005), Tree allometry and improved estimation of carbon stocks and balance in tropical forests, *Oecologia*, *145*, 87.

Chave, J., H. C. Muller-Landau, T. R. Baker, T. A. Easdale, H. T. Steege, and C. O. Webb (2006), Regional and phylogenetic variation of wood density across 2456 neotropical tree species, *Ecol. Appl.*, *16*(6), 2356.

Drake, J. B., R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdone, J. F. Weishampel, and S. D. Prince (2002), Estimation of tropical forest structural characteristics using large-footprint lidar, *Remote Sens. Environ.*, *79*, 305.

Dubayah, R. O., S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R. L. Chazdon (2010), Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica, *J. Geophys. Res.*, *115*, G00E09, doi:10.1029/2009JG000933.

Hajnsek, I., F. Kugler, S. K. Lee, and K. P. Papathanassiou (2009) Tropical forest parameter estimation by means of Pol-InSAR: The INDREX-II Campaign, *IEEE Trans. Geosci. Remote Sens.*, *47*(2), 442.

Hamilton, W. C. (1964), *Statistics in Physical Science*, Ronald, New York.

Houghton, R. A., F. Hall, and S. J. Goetz (2009), Importance of biomass in the global carbon cycle, *J. Geophys. Res.*, *114*, G00E03, doi:10.1029/2009JG000935.

Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding (2002), Lidar remote sensing for ecosystem studies, *BioScience*, *52*(1), 19.

Lu, D. (2006), The potential and challenge of remote sensing-based biomass estimation, *Remote Sens. Environ.*, *27*(7), 1297.

Rodriguez, E., and J. M. Martin (1992), Theory and design of interferometric synthetic aperture radars, *IEE Proc.*, *139*(2), 147.

Treuhaft, R. N., S. N. Madsen, M. Moghaddam, and J. J. van Zyl (1996), Vegetation characteristics and underlying topography from interferometric radar, *Radio Sci.*, *31*, 1449–1485, doi:10.1029/96RS01763.

Treuhaft, R. N., G. P. Asner, and B. E. Law (2003), Structure-based forest biomass from fusion of radar and hyperspectral observations, *Geophys. Res. Lett.*, *30*(9), 1472, doi:10.1029/2002GL016857.

Treuhaft, R. N., B. D. Chapman, J. R. dos Santos, F. G. Gonçalves, L. V. Dutra, P. M. L. A. Graça, and J. B. Drake (2009), Vegetation profiles in tropical forests from multibaseline interferometric synthetic aperture radar, field, and lidar measurements, *J. Geophys. Res.*, *114*, D23110, doi:10.1029/2008JD011674.

B. D. Chapman, Jet Propulsion Laboratory, California Institute of Technology, MS 300-235, Pasadena, CA 91109, USA. (bruce.chapman@jpl.nasa.gov)

J. R. dos Santos and L. V. Dutra, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP 12227, Brazil. (jroberto@dsr.inpe.br; dutra@dpi.inpe.br)

J. B. Drake, USDA Forest Service, Tallahassee, FL 32303, USA. (jasondrake@fs.fed.us)

F. G. Gonçalves, Jet Propulsion Laboratory, California Institute of Technology, MS 138-307, Pasadena, CA 91109, USA. (fabio.goncalves@jpl.nasa.gov)

P. M. L. A. Graça, Instituto Nacional de Pesquisas da Amazônia, Manaus, AM 69060, Brazil. (pmlag@inpa.gov.br)

G. H. Purcell, Jet Propulsion Laboratory, California Institute of Technology, MS 138-308, Pasadena, CA 91109, USA. (george.purcell@jpl.nasa.gov)

R. N. Treuhaft, Jet Propulsion Laboratory, California Institute of Technology, MS 138-212, Pasadena, CA 91109, USA. (robert.treuhaft@jpl.nasa.gov)