There is substantial uncertainty in our knowledge of existing forest carbon stocks and their spatial distribution [1]. Moreover, the effects of human induced changes on the terrestrial carbon cycle through processes such as deforestation and associated regrowth are also uncertain, making it difficult to quantify the exchange of carbon between the surface and the atmosphere, and therefore to predict the consequences of such change on atmospheric CO₂ [2].

Given the rapidity of vegetation structure changes — whether through mortality, deforestation, disturbance, regrowth — and the vastness of the Earth’s forest resources, in situ monitoring is essential, but cannot feasibly be accomplished at spatial resolutions that are commensurate with the changes that are taking place (which typically occur at scales of 1 ha or below) [3]. Given these factors and their influence on uncertainties in the global exchange of carbon between the land surface (particularly forests) and the atmosphere, satellite observations are required to not only better map the extent for vegetation types, but also to provide better estimates of forest carbon stocks, especially aboveground biomass (AGB).

Aboveground biomass is particularly relevant for a number of reasons:

- It is most susceptible to change through both natural processes (e.g., fire and insect disturbance) and human activities (e.g., deforestation and forest degradation);
- It is often used as the basis to estimate other terrestrial carbon pools (e.g., litter, dead wood and belowground biomass);
- It provides an indication, when used with other forest and environmental variables (e.g., age and elevation), of potential for carbon sequestration from the atmosphere to additional biomass accumulation (i.e., net primary production).

In this article, we discuss the potential of new technologies to advance the estimation of AGB at high resolution over large areas, even globally, which in turn has the potential to substantially reduce uncertainty in global carbon exchanges and net carbon budgets and, thus, to improve our knowledge of the magnitude and net changes...
in various terrestrial carbon pools (including soil carbon). We do not suggest that satellite observations are the sole answer to resolving uncertainty in global carbon budgets, nor do we believe that they will obviate the need for in situ measurements. On the contrary, field measurements of forest carbon stocks are essential for both calibrating and validating (i.e., assessing) spatial estimates of AGB derived from satellite observations. This is particularly true when the satellite and field data are linked across scales using datasets judiciously acquired from aircraft platforms.

Before we present our perspectives on the utility of new remote sensing technologies to inform and improve international policy initiatives, such as those of the United Nations Framework Convention on Climate Change (UNFCCC), we first provide a brief overview of the current state of carbon stock and emissions reporting requirements under the evolving Reducing Emissions from Deforestation and forest Degradation (REDD) and REDD+ framework (described below). We place this discussion in the context of national carbon emissions reporting and associated evolution of current ‘measuring, reporting and verification’ (MRV) requirements.

Reducing emissions from deforestation & forest degradation

The UNFCCC has developed a framework for addressing carbon emissions from land-use change, particularly deforestation, having recognized that 10–20% of global greenhouse gas (GHG) emissions originate from these activities [4,5]. As part of recent international agreements under the UNFCCC, and specifically REDD, the Subsidiary Body on Scientific and Technological Advice (SBSTA) has been tasked with developing modalities relating to, “Measuring, reporting, and verifying anthropogenic forest-related emissions by sources and removals by sinks, forest carbon stocks and forest area changes resulting from the implementation of REDD+ activities, consistent with any guidance for MRV [Annex II (c)].”

This is associated with the development of robust and transparent national forest monitoring systems (NFMS) for the monitoring and reporting of REDD+ activities, tailored as necessary to national circumstances, which we note are highly variable with respect to technological capacity. Notably, paragraph 71(c) of Decision 4/CP.15 (agreed at the Conference of the Parties in Copenhagen), recognizes the “…establishment of robust and transparent national forest monitoring systems and, if appropriate, subnational systems as part of national monitoring systems, that use a combination of remote sensing and ground-based forest carbon inventory approaches for estimating, as appropriate, anthropogenic forest-related GHG emissions by sources and removals by sinks, forest carbon stocks and forest area changes.”

The inclusion of subnational systems reflects the reality that REDD+ activities are advancing within national borders (e.g., a compliance program linking forest protection at the state level in the Brazilian Amazon with the California Climate Initiative to offset emissions). Decision 4/CP.15 also stipulates that any NFMS provide estimates that are transparent, consistent, as accurate as possible, and those that reduce uncertainties, taking into account national capabilities and capacities. The outcomes of NFMS must be widely available and suitable for review by interested parties and experts, whether from within or outside the reporting nation.

Measuring, reporting & verification

The measuring component of MRV consists of documenting the extent and associated changes in forest area and the carbon stock associated with those changes. These changes need to be attributed to causes; for example, so called ‘activity data,’ which in turn implies mapping of land use and management practices. It also requires information on so-called ‘emission factors,’ which describe the carbon content in various terrestrial pools (i.e., AGB, belowground biomass, dead wood, litter and soil organic matter). A framework for reporting emissions in this manner is provided, in overview, by Maniatis and Mollicone [6]. For the purposes of MRV, measurement over time implies monitoring, since change is an implicit reporting requirement under the UNFCCC, and forms the basis for any financial compensation under either market-based or voluntary emissions-reduction schemes [7]. Again, measuring land-use activities and carbon pool emission factors requires in situ measurements, and both of these can be informed by remote sensing observations. This approach forms the basis for international initiatives, such as those of the Group on Earth Observations [8]. Currently, such observations are essential for identifying activity data, and in this article we argue that they also can substantially improve the emission factor datasets.
used on a national basis. In particular, extending field measurements requires some sort of spatial extrapolation, whether done through sampling methods (the traditional approach) or via more extensive sampling (i.e., so-called ‘wall-to-wall’ mapping) [9]. We note that even the traditional sampling methods benefit from making use of some sort of satellite data; for example, for land cover or land use classification and stratification. Both land cover and land use mapping require in situ information to produce accurate maps, and, thus, attribution (activity) data, but this is particularly true in the case of land-use mapping, because satellites can only observe biophysical attributes of the surface, not specific human uses of the land surface [10].

We focus here primarily on measuring and monitoring (M&M), and mapping (as a result of our focus on wall to wall satellite observations). Reporting carbon stocks and emissions is largely the responsibility of the countries that choose to participate in the REDD+ mechanisms, for sovereignty reasons and, thus, we only note here that the large disparity in technical capabilities requires a concerted effort for nationally focused technical-capacity building (and this has been the focus of a number of international development efforts, particularly the UN’s REDD program). Similarly, the verification component of MRV will probably be partly conducted via a review phase by the UNFCCC and partly by third-party organizations. Thus, we do not focus on it here other than to acknowledge that many of the same approaches we discuss under M&M are applicable to verification, particularly those related to scaling from in situ measurements to large areas using aircraft-based observations over intermediate spatial domains. There is a possibility that Annex-I nations that invest substantial financial resources into supporting developing countries’ REDD+ participation may choose or be mandated by their respective governments to provide their own verification ability (e.g., in the case of the USA, an agency such as NASA may be tasked to support this process). We note that the verification process is crucial in that it influences the quality reporting tier [6], which in turn, determines the financial value of any carbon credits or compensated emission reductions.

M&M informed by satellite observations

There are a number of ways in which satellite observations can inform national emission reporting. As noted earlier, satellite data are now routinely used for classifying land cover types and can also be used to infer land-use information [9]. Whereas there have been many land cover/land use (LC/LU) classification schemes developed over the years [11], for the purposes of reporting under a REDD+ mechanism, schemes that permit activity data reporting under UNFCCC guidelines are the most relevant (i.e., forest land, cropland, grassland, wetland, settlement and other land). Moreover, these categories can be used to stratify in situ sampling such that it better represents the spatial extent of various types of land cover and land use. This is also true with broad categories of land cover (e.g., forested lands), where additional definition of types and densities of forest cover can be further refined, for example, forest inventory efforts, such as those spearheaded by a wide range of countries under the United Nations Food and Agricultural Organization Forestry programs [101]. However, it is a challenge for inventory approaches to capture the small grain size of disturbance over many areas (e.g., in Africa where most change is associated with small-land holders). Many of these same areas also have outdated inventories. Remote sensing can aid this by identifying such areas for more effective sampling. This approach provides a basis for Tier-1 reporting under the UNFCCC and is particularly effective at that level when sampling efforts are focused not only on areas of change, but also include adequate sampling of the large areas where change may be effectively quite small (e.g., forest that remains forest) but above-ground carbon stock both large and spatially variable.

Another way in which remote sensing can inform carbon stocks and emissions reporting is by extending in situ measurements over larger areas by flying aircraft equipped with various sensors over sites where intensive inventory data have been collected. This approach has been used to provide estimates at the landscape scale, when combined with LC/LU classification schemes, using both active remote sensing, in particular airborne and satellite-based light detection and ranging (LiDAR) [12-14], but also other remote sensing datasets [15,16]. While this approach is advantageous in terms of better characterizing fine scale spatial variability in terrestrial carbon stocks across large spatial extents, it is ultimately limited in that aircraft data are not an ideal method to routinely and repeatedly acquire systematic observations over very large (i.e., continental to global) spatial extents.

A closely related approach is to use airborne or satellite LiDAR to inform ecosystem models, whether statistically or mechanistically based, using a combination of in situ measurements with the remotely sensed observations. This has been accomplished across a range of spatial scales using a wide range of models [17], but perhaps most effectively using models that are intrinsically based on allometric relationships between various forest structure metrics and AGB, such as those used in the Ecosystem Demography model [18]. Ecosystem Demography model is particularly suited for use with LiDAR data because it is height-structured, so that initialization of mean canopy height initializes
both above- and belowground states and provides direct information on successional state (under the assumption that short canopies are young).

A final approach that we consider here is one based on the direct relationships between biophysical attributes of vegetation and remote sensing observations acquired either by aircraft or satellite sensors. The utility of this approach relative to more traditional methods to estimate AGB has been described elsewhere [19]. Essentially, it is possible to develop robust relationships between forest attributes, such as canopy height, basal area and diameter at breast height (DBH) and, for example, radar backscatter at different polarizations or various LiDAR metrics (such as the difference between the initiation and the cessation of the laser pulse or the characteristic height at which half of the sensor-transmitted energy is returned). This latter topic is explored in substantially more detail below. Our contention is that a direct remote sensing approach, informed by field data and basic ecological principles, has the potential to substantially advance national to global scale reporting of carbon stocks and emissions. An example of one such map recently derived using this approach is shown in Figure 1.

**Challenges of M&M with remote sensing**

The use of emission factors is largely based on in situ measurements within various terrestrial pools, and yet these are often woefully inadequate for capturing the range and spatial variability within and between the various pools. Satellite remote sensing also faces challenges when applied to carbon stock estimation or monitoring change through time. There are, for example, limitations of remote sensing maps and associated emissions-reporting schemes that depend intrinsically on LC/LU classifications, since these are notoriously prone to error (misclassification), despite a range of recent advances in methodological techniques, unless the number of LC/LU classes is kept to a relatively small number (e.g., 10) and rich in situ data

---

**Figure 1.** Map of aboveground biomass across the Brazilian Amazon derived from systematic global-scale satellite observations. The inset provides an indication of the spatial detail inherent in the map, which was produced using a combination of geoscience laser altimetry system (GLAS) light detection and ranging observations and moderate resolution imaging spectroradiometer image datasets. Image courtesy of Baccini A; the Woods Hole Research Center, MA, USA.
are available to constrain the possible outcomes [20]. This is particularly true in the case of change detection and systematic monitoring of those changes through time, partly because errors are magnified as additional time steps are added, unless a time series of images are used to detect change retrospectively [21]. For these reasons, among others, it is essential to report error and uncertainty in any carbon stock or emissions estimates, including those based on activity data derived from LC/LU maps.

Similarly, the application of remote sensing techniques to estimate AGB, which would aid extension of the limited emission factors to larger areas, have been limited across the ranges of biomass encountered, and with inadequate accuracy to meaningfully monitor and manage carbon stocks. For example, optical imagery may not be sufficiently sensitive to variations in AGB unless these are associated with variations in upper canopy structure, shadowing and photosynthetic capacity (which they often are). Frequent multi-temporal observations such as those from the moderate resolution imaging spectroradiometer help in this regard, and newer sensors that make use of multi-angle data and possibly hyperspectral measurements may push this further forward as well. Even radar-based estimates of AGB at the stand level have consistently shown saturation levels at levels exceeding 100 Mg ha\(^{-1}\) (50 Mg C ha\(^{-1}\)). This limits the utility of radar for direct estimation of AGB over much of the world’s tropical forest areas when radar is used as the sole source of remotely sensed observations. Although some individual studies have pushed this upper limit, (e.g., with P-band or VHF systems) there has been no definitive demonstration that existing radar techniques can routinely exceed this amount. This may change with the further development of multi-band and multi-polarization radar techniques, especially ones that include some limited LiDAR data. In addition, polarized interferometric synthetic aperture radar (pOLInSAR) holds great promise for deriving structure and biomass beyond what is achievable today [22].

Despite the limitations described above, several of which also affect inventory-based approaches, we contend that new remote-sensing technologies are more than adequate for directly estimating AGB, are at least comparable to those based solely on inventory data alone, and that an optimal approach would make use of the integration of both direct and inventory-based approaches. Whatever the techniques employed, once AGB is estimated (with appropriate uncertainty metrics), it can be used to augment the estimation of other terrestrial carbon pools (and this is often the basis for traditional methods as well, for example, using AGB to estimate litterfall, dead wood and belowground biomass). This perspective cannot be simply dismissed as academic or irrelevant to REDD+ reporting, not only because AGB is useful for estimating other C pools but also because it can substantially augment inventory-based approaches to estimate AGB (as further discussed below).

Over the next decade, LiDAR remote sensing, particularly when combined with other data sources such as InSAR, will dominate efforts to map and monitor biomass at local to regional scales, and for this reason we will focus the remainder of our discussion on the use of LiDAR. One disadvantage of LiDAR, as with all optical systems, is that it cannot make accurate measurements through clouds (as opposed to radar). Thus, areas that are persistently cloudy may be under-sampled, such as in many tropical areas (e.g., over Gabon). However, repeated frequent orbits gradually fill in even very cloudy areas and, unlike passive optical sensors, LiDAR is active, providing its own illumination and, thus, makes measurements at night (when convective cloud cover in tropical regions is reduced).

A vegetation LiDAR survey from space will provide a definitive baseline for carbon stocks and, while it may not be the most desirable solution, is reasonable to assume that only one such baseline may be required (e.g., over a decade) and that other operational satellites, such as radar or multispectral optical, may be used to monitor changes through time. These operational satellites could be tasked to monitor deforestation and, with suitable algorithm development, degradation [9]. A LiDAR survey that provides height, along with some estimate of age, such as from a Landsat disturbance product, then enables the tracking of natural regrowth.Moreover, a one-time LiDAR survey provides the basis for continuous updating via in situ field sampling. A stratified sampling of LiDAR footprints, which are precisely located, could be visited in the field and re-measured every few years (similar to other inventory plots). In this case, the LiDAR data provide the core knowledge needed to created an unbiased sampling network at national to continental scales. Yet a third advantage of space-based LiDAR is that it provides enough examples of forest structure across the variety of forest types and gradients to facilitate the development of remote sensing algorithms that are not based on LiDAR (i.e., those based on radar or multispectral optical).

We feel compelled at this stage to reiterate two critical points:

- As should be clear from the discussion thus far, we are in no way de-emphasizing the importance of in situ measurements and inventory-based approaches, rather we view them as essential datasets that can be
used synergistically with remotely sensed observations to improve what can be accomplished with one or the other alone;

- Similarly, we are not dismissing radar (or, for that matter, optical remote sensing) approaches to mapping AGB, rather we view the data from such instruments as critical components of an effective forest monitoring system because they can be used to extend essential LiDAR measurements to much larger spatial domains, as described below. Furthermore, the continued development of multi-band and multi-polarization techniques, especially ones that include some limited LiDAR data along with the advent of polarized interferometric synthetic aperture radar (polInSAR) holds great promise, but is beyond the scope of this article to discuss this at length here. Suffice it to say that the foreseeable paradigm of radar being used to extend LiDAR observations may be flipped on its head where very limited LiDAR data are used solely to calibrate spatially continuous polInSAR data for mapping height, vertical structure and AGB.

**Opportunities for improved M&M**

The use of LiDAR remote sensing to map vegetation properties has grown dramatically in the last decade [23], since early forestry applications in Sweden [24] and Finland [25]. LiDAR is now routinely used to measure a large array of canopy properties with an accuracy and efficacy that exceeds, often greatly, what other passive and active remote sensing instruments can accomplish. The data that LiDAR provides on the vertical variation of canopy elements is so unique that it has the potential to not only augment traditional ground-based methods, but to open up wholly new research avenues in ecosystem science in areas such as biodiversity and habitat characterization [26] and forest fire fuels modeling [27], in addition to biomass mapping.

By far, the most common source of LiDAR data has been from airborne, commercial small-footprint LiDAR systems. Originally sought for their ability to produce remarkably detailed, fine scale topographic information, these systems are increasingly applied to forestry applications. However, because they image relatively small areas, application to large-scale monitoring of forests may be prohibitively expensive without some sort of sampling scheme similar to those described for inventory-based methods discussed earlier [12]. Moreover, they do not form the basis of space-based systems. Rather, it is a large-footprint, waveform systems that provide the foundation for global monitoring and inventory of vegetation structure from space.

Airborne analogues for space-based LiDAR data, such as the laser vegetation imaging system (LVIS) [28], have been useful for exploring the potential of vegetation LiDAR on satellite missions. LVIS is a large-footprint waveform based instrument that has served as a primary validation and research tool for the application of LiDAR to ecosystem studies. LVIS has imaged millions of hectares across numerous biomes throughout the world and these data have formed part of the basis of our understanding on the use of LiDAR to derive vegetation structural characteristics.

In the next decade, space-based characterization of the global distribution of AGB, as well as terrestrial sources of carbon resulting from disturbance and recovery, will be provided by a combination of satellite missions. One such, the deformation, ecosystem structure and dynamics of ice (DESDynI) mission [29] was essentially cancelled by NASA, probably owing to budget (but not science) limitations [30]. DESDynI was to be comprised of two instruments, a multi-beam LiDAR and an interferometric SAR (NASA continues to study the SAR element of this mission, but the LiDAR instrument was terminated with the hopes that an international collaborator may be able to contribute that segment). The science objectives of DESDynI, which will be applicable to other descendant missions, required global mapping at scales as fine as 1 ha to capture the characteristic grain size of disturbance events [3] including, for example, forest degradation associated with selective logging operations [31,32] and structure changes and recovery following fire events [33,34]. These future missions will rely on the synergistic use of LiDAR with radar and optical imagery to accomplish spatially complete and contiguous mapping. Although not designed for vegetation monitoring, the ICESAT2 mission (scheduled for launch after 2016) will provide some capability for mapping canopy heights in forests with cover that do not exceed approximately 70% [35]. Plans also are under development to scope a possible vegetation LiDAR mission using the international space station.

There have been several review papers that have examined the utility of LiDAR for retrieving forest and vegetation structural measurements for ecosystem science [23,36]. The literature on the use of small-footprint LiDAR for biomass estimation is large and growing exponentially. Our goal here is to briefly summarize what we know about the accuracy and efficacy of LiDAR with a specific focus on space-based retrievals (and their analogues) of forest AGB. Ultimately, we believe it will be satellite-based methods that provide the stable, consistent, accurate and transparent platform required for REDD+ and other forest monitoring frameworks. We begin by first briefly reviewing the fundamentals of waveform LiDAR remote sensing.
LiDAR fundamentals
There are many types of LiDAR systems that differ with regards to footprint size (spatial resolution), wavelength, pulse rate, scanning (sampling) pattern and return signal digitization. The choice of, and sometimes the trade-offs between these variables have implications for the retrieval of biomass. For systems observing vegetation the wavelength is generally in the near-IR (e.g., 1064 nm) where canopy elements are highly reflective and solar background noise is limited. With most systems, a pulse of laser energy is emitted towards the surface. This energy then reflects off various elements of the surface, such as leaves, branches and ground. The roundtrip return-time of the pulse is recorded, providing a range or distance to the target (and hence the ‘ranging’ part of the LiDAR acronym). In the case of a vegetated surface, the return pulse is attenuated, with energy returned in proportion to the amount and reflectance of material at various heights as the emitted pulse travels through the canopy and reflects off the ground (Figure 2). This is what is meant by a ‘waveform.’ Space-based LiDAR cannot have high pulse rates (nor small footprint sizes) as the power required to operate them is too great. Orbital configuration is another consideration for LiDAR space missions since it relates not only to geolocation, but also spatial sampling and completeness (coverage). Space-based LiDAR geolocation accuracy, such as those estimated in the design of the DESDynI mission, is approximately 5–8 m (horizontal); thus, coordinated in situ measurement plots need to be large enough to encompass potential footprint location uncertainty to ensure unsampled elements do not bias estimates. Orbits can be designed such that missed spaces between orbits are filled in by successive passes. For this reason it is often useful to speak about the final number of ‘shots’ per unit area an instrument will achieve. The DESDynI mission

Figure 2. The major components of a return light detection and ranging waveform. The waveform returns are converted from time to distance and positioned in elevation space relative to the ellipsoid of the Earth (the x-axis on the left). The ground is located by finding the center of the last return (shown in maroon). Once the ground is found the top of the canopy expressed as a height is computed by subtracting the first return (called the ‘highest reflecting surface’) from the ground return. Various waveform metrics may be calculated from both the waveform and the cumulative return energy profile, which is found by summing the energy from the ground return and canopy return to the top of the canopy. The quantiles of the cumulative energy profile are often used for biomass estimation, such as RH50, which is the height of the median energy return. Note the highest reflecting surface may not correspond precisely to the top of the canopy because the beam may penetrate slightly into the canopy before enough leaf material is intercepted to provide a strong enough return.
scoping planned for approximately 50 observations in a 1 km cell at the equator after 2 years (with denser coverage away from the equator towards the poles). Other configurations, for example, the international space station, could achieve substantially more observations depending on how many beams are deployed and the ultimate mission parameters. This emphasizes the sampling nature of space-based LiDAR and why synergistic use with imaging sensors, such as optical and radar systems, is required for a complete mapping of AGB in the foreseeable future. That said, the next generation of LiDAR instruments may very well have true imaging capability and studies are underway to develop such technology (e.g., by providing wall to wall footprints in swath widths of several kilometers).

The basic measurement from LiDAR is an amplitude (reflectance) as a function of distance (height), as represented by the waveform. The amplitude of the waveform at a particular height is a function of how many photons are received by a detector that appear to originate from that height. The number of photons in turn is determined by the reflectance of the material, their orientation, and the attenuation of the incoming beam through the canopy. In the case of vegetation the waveform is essentially a record of the vertical distribution of leaves and their supporting branches: where the amplitude is larger, there is more canopy material, and where it is smaller there is less.

The area under the waveform curve represents the total number of photons received or total energy. The amplitudes may also, thus, be scaled to this total energy. This facilitates comparisons between waveforms. We call this ‘scaled waveform’ (again meaning the amplitude is scaled to the total energy in the waveform) [37]. The waveform is used to create other LiDAR metrics that are either directly or indirectly used to find the canopy variables of interest.

Some of the essential features of the waveform are highlighted in Figure 2. The ground return is that portion of the waveform that originates from reflectance off the ground underneath the canopy. The elevation of the ground is found by determining the center of the ground return portion. The canopy return is that portion of the waveform that originates from the canopy components themselves. The top of the canopy is found by starting at the ‘leading edge’ of the return until a signal greater than some noise threshold is found. It is sometimes impossible to distinguish the boundary between the two returns (canopy and ground) because of understory and slope effects that blur the ground return [38].

- **LiDAR canopy metrics**

  The maximum canopy height is found from the top of the canopy return in the waveform, or may be some percentile down from it (e.g., at 98% of the cumulative energy return). Although we call this ‘canopy height’ it is actually the highest reflecting surface. There is always some penetration of laser energy through the canopy before enough energy is returned for detection. Other canopy height metrics are often created for other energy quantiles that give the height at which that energy from the return is reached (e.g., [38–40]). It is worthy of note that small-footprint data can be aggregated to form ‘pseudo-waveforms’ to create similar metrics.

  Other variables useful for forest monitoring and biomass estimation can be derived from the waveform data, such as total and vertical canopy cover, LAI and transmittance [37]. In addition, many other metrics are directly derivable from the waveform, and generally have to do with statistics on the vertical variation within waveforms, or the vertical and spatial variation among waveforms. Lefsky et al. summarize and review various parameters that are derivable from LiDAR waveforms, but they found a high degree of redundancy (i.e., although many different types of parameters can be created from waveforms, they often contain redundant information) [41].

  Canopy height is arguably the most common and desired canopy variable derived from LiDAR for ecosystem studies. However, it is also one of the most difficult to validate because of the vagaries of ground measurement and definitions of ‘height’ over extended footprints. Some typical height measures include crown-weighted height (a surrogate for Lorey’s height), the mean of the maximum heights in an area, and amplitude weighted height. A common height measure used with small-footprint data is the mean height of all canopy returns [42].

  Although there have been many studies validating the retrieval of canopy heights from small-footprint LiDAR, there have been considerably fewer using large footprint LiDAR. In general, 3–4 m errors can be expected at medium (~30 m) footprints. For larger footprints such as from ICESAT (with diameters exceeding 65 m), errors will increase, especially over steep slopes.

  Dubayah et al. suggest that a significant source of error is in the ‘round truth’ itself since accurate in situ height measurements are hard to achieve in dense, closed-canopy forests [38]. Repeated ground measurements by different field crews can lead to regression root mean square error (RMSE) errors of 2–3 m in our experience in the tropics [43]. By contrast, the instrument error from planned missions is approximately 1 m [29,38]. While instruments operating from space may theoretically be able achieve such high accuracies, in practice, the effects of geolocation error, slope, combined with errors in field sampling may result in errors at the footprint scale that are closer to what has been observed from aircraft studies. However, it is critical to note that
Carbon stock retrieval methods & accuracies using LiDAR

No other biophysical variable has received as much interest for derivation from LiDAR than biomass. Indeed, estimating the carbon content of forests and its dynamics is the primary driving force behind the development of space-based LiDAR missions. There have been numerous applications of large footprint LiDAR to estimate biomass [38, 40, 44, 45]. Such studies have shown accuracies (RMS error) that have ranged from approximately 20 to 200 Mg/ha at plot scales of generally approximately 30 m (900 m²) but as large as 1 ha. Relative to mean biomass levels, the errors in these stand-level studies have a mean of approximately 20%. Results from small-footprint studies are generally in the same range.

One confounding factor in evaluating LiDAR-based efficacy is that different ‘accuracies’ are routinely reported by investigators. An RMSE is commonly reported, but this can be overly optimistic of future accuracies. Cross-validated RMSE magnitudes are far more valuable and indicative of true accuracy. At the mapping or pixel level, very few studies provide a spatial map of errors or confidence intervals for specific pixels predicted using a regression equation (e.g., [38]). Lastly, it is important to understand that domain-wide accuracies, say over a region, will always be less than the accuracies achieved at finer scales because in the absence of bias, the error will be reduced as more observations are used to estimate the mean (after accounting for spatial autocorrelation exactly as described above for heights). As an example, Gonzalez et al. report small-footprint LiDAR accuracies of <1% for the mean regional biomass, but the mean of individual pixel errors at 30 m resolution exceeded 44% [46].

The basis of estimating biomass from LiDAR comes primarily from the relationship between some type of measure of height or canopy structure and biomass. In its simplest form, taller trees weigh more then smaller trees, and this relationship seems more then smaller trees, and this relationship seems to hold for assemblages of trees and across species. Waveform LiDAR is rarely used to estimate the biomass of an individual tree; rather, there are generally several, sometimes tens to hundreds of individuals, in a footprint. Thus, studies with large footprint waveform LiDAR are concerned with estimated plot-level biomass and beyond. Generally, this is carried out through statistical methods, most commonly, multiple-regression or more sophisticated machine learning techniques, such as regression trees. Ground data are collected, either within a LiDAR footprint, or for an area that encompasses several LiDAR footprints (e.g., a 1 ha plot). Biomass is found allometrically, generally at the species level, using DBH alone, or where available, and generally more accurately, DBH and height (e.g., after [47]).

Linear equations relating maximum canopy height from LiDAR, by itself, to total AGB are often not as accurate as equations that use some non-linear relationship and use other LiDAR height metrics (but this is a function of species assemblages such as deciduous vs coniferous). Non-linear relationships with height itself are not surprising given the non-linear relationship exhibited by many species of biomass with height. In addition, the relationship between DBH and height for a species is often non-linear itself as age increases, so that the degree to which height is used as surrogate for DBH in biomass prediction may affect the form of the equations found.

However, as we have mentioned, ‘height’ from LiDAR can have several meanings, especially when applied across several footprints to make an estimate at the plot level. It may mean the average maximum canopy height, it may be the mean canopy height of all material in the waveform, or it may be any number of related metrics, such as height², quadratic mean canopy height and so on. All of these have been used in estimation of biomass at the plot level.

More fundamentally, while height from large footprint LiDAR can tell us ‘how tall’, it does not tell us ‘how many.’ For open canopy forests, methods using either small-footprint LiDAR data or high-resolution optical data are often used to get tree densities. For large footprint data, other approaches may be required that go beyond site-specific statistical relationships that are not easily derivable and not transferable across domains. There are perhaps ecological scaling features of canopy growth and dynamics that relate the height of a canopy to its biomass and stem density; in that, when canopies reach a certain height under various edaphic and climatological conditions, stem density may vary in a predictable way [48]. This is the basis for gap-based models such as the Ecosystem Demography model that successfully use average maximum canopy height from LiDAR (at 1 ha) to initialize model state for biomass stock and flux prediction [18]. However, scaling theory may be used directly with LiDAR and radar data to constrain estimates of stem density: radar would provide estimates of volume, and LiDAR heights, which together may be sufficient to estimate stem density as constrained by allometry.
The exact mechanisms that explain why LiDAR canopy height metrics are such good predictors of plot-level biomass is still a subject of research. Canopy height may not necessarily be a good predictor of crown biomass, but rather a more sensitive predictor of total stand (including stem) biomass. However, at the plot level, the ‘how many’ question (i.e., stem density and related basal area) is important and, for this reason, other height metrics have been successfully used to improve biomass predictions. These metrics provide information about how open and closed the canopy is, the strength of the ground return, and the distribution of the vertical material, which in turn seem to be a function of stem density.

Based on these observations and associated scale considerations, we conclude that LiDAR can meet a biomass requirement of 20 Mg/ha or 20%, assuming there are sufficient numbers of observations per hectare. This is not an issue for aircraft LiDAR, but to take REDD+ MRV to scale requires repeated and consistent measurements that we believe can best be accomplished using space-based LiDAR. Currently, scoped space-based vegetation LiDAR missions would probably not achieve sufficiently large numbers of LiDAR shots to reliably estimate biomass at grid-cell sizes of less than approximately 0.5 km. Thus, the need for synergistic use (often called fusion) with other more spatially complete datasets (whether radar, optical or both), is underscored. This approach for direct AGB mapping and monitoring obviates the need for land use or cover type change classification, although that could still be incorporated post hoc to ascribe causes (i.e., activity data).

### Carbon stock dynamics

One of the most exciting uses of LiDAR is to monitor the dynamics of forest carbon stock changes through time. This is what is known in UNFCCC parlance as a ‘stock difference method.’ One of the causes of the large uncertainty about the emissions of carbon from the land surface come from the inability to correctly identify and quantify changes in sources and sinks. LiDAR can be used to estimate carbon stocks, and thus, changes in these stocks through subsequent deforestation. The area deforested can be found through multispectral imagery, such as Landsat and LiDAR data can be used to place a biomass value at a particular point in the chronosequence. However, estimating changes from increases in growth, from natural disturbance and mortality events, and recovery from these disturbances from multispectral or radar imagery is much more difficult. Thus, there is considerable interest to evaluate the ability of LiDAR to capture rates of regrowth and degradation. The ability to monitor disturbance, degradation and regrowth between two points in time from space would essentially determine and quantify which areas are sinks or sources of carbon and, furthermore, would directly and immediately yield the net terrestrial flux of carbon between the atmosphere and terrestrial vegetation. Dubayah et al. created a map of sources and sinks of carbon for a tropical forest using this approach. Figure 3 shows an example for the tropical forests at the La Selva Biological Station in Costa Rica.

We highlight that from a variety of perspectives it is the change in biomass (i.e., the carbon stock change) and the sequestration potential of an area that is of most interest. Simply knowing a forest’s biomass and its height is not sufficient: a short stature forest may be short, not because it is young, but because it is site-limited (e.g., by soils or elevation). Repeat LiDAR measurements of height may allow us to observe growth increments and, when combined with height, an inference of age and, thus, sequestration potential.

### Future Perspective

Over the next 5 years, measurement of AGB will be dominated by methods that combine sparse and airborne LiDAR with spatially continuous data such as from passive optical and radar. While there will be a great deal of improvement in methods, the lack of LiDAR data over large areas and the expense of obtaining it repeatedly and consistently will continue to be the fundamental limitation. One new technological development that might ease some of this burden is ‘photon counting,’ a LiDAR technology that measures the proportion of photons reflected from a surface. Photon counting, if it can be scaled to high-altitude aircraft, would be able to provide wide-area coverage and fine resolution simultaneously, and may be a significant advance over existing commercial systems in terms of spatial coverage. However, this is nascent technology and much uncertainty surrounds its development and ultimate efficacy for forest mapping.

As we look beyond 5 years, we enter the era of space-based vegetation LiDAR missions that will revolutionize the estimation of carbon stocks at regional, national and global scales. LiDAR, if deployed with current radar technology as well as other remote sensing data sources, could map biomass at spatial resolutions of 100 – 250 m globally (and indeed these were the spatial requirements for the DESDynI mission). It is thus reasonable to project that such measurements from space will be used to form the first consistent and transparent baseline of global carbon stocks at policy
relevant (i.e., both national and subnational) spatial scales. Finally, there is a mission concept using polInSAR that could take advantage of two sister spacecrafts that contain duplicate L-band InSAR sensors. The two instruments would observe the forest canopy at the same time, but from different angles. There is evidence that simultaneous observations of forest canopy using InSAR, calibrated with LiDAR data, could provide not just height, but vertical canopy profiles and biomass globally at resolutions <100 m. If a tandem polInSAR mission potential is realized, it would represent an even greater leap forward for the measurement and monitoring of forest structure and biomass globally. Our own view beyond 10 years is that carbon monitoring will be dominated by a fusion of sensor data that relies heavily on LiDAR and SAR/polarimetric InSAR, both from aircraft and from space.

**Acknowledgements**

We acknowledge the anonymous reviewers for their constructive comments, as well as the funders for their support. We thank Alessandro Baccini, Skee Houghton, Nadine Laporte, Mindy Sun and Anu Swatantran for fielding questions and assistance with graphics.

**Financial & competing interests disclosure**

The work reported here was supported by the US National Aeronautics and Space Administration (NASA) Carbon Cycle and Terrestrial Ecosystems program (grant NNX09AI24G to Scott Goetz and NNX08AP55G to RD). Scott Goetz also acknowledges coordinated support from the Gordon and Betty Moore Foundation, the Google.org Foundation and the David and Lucile Packard Foundation. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed. No writing assistance was utilized in the production of this manuscript.

**Figure 3.** Changes in height and biomass over tropical forests as measured using aircraft-acquired laser vegetation imaging system data in 1999 and again in 2008 at La Selva Biological Station in Costa Rica. Changes in AGB between dates were calculated at 30 m resolution, then statistically assessed for significance at the 1 ha level. The colored grid shown gives the changes in AGB (with the underlying height of the forest canopy shown below it). The inset shows the changes mapped over a shaded perspective view of canopy structure obtained from small-footprint LiDAR (not used in the AGB estimates). Such change data allow for the mapping of carbon sinks and sources across the landscape, but must be evaluated carefully for errors to determine significance (see for example Dubayah et al. [37]).

AGB: Aboveground biomass.
Executive summary

Measuring % monitoring informed by satellite observations
- Satellite and aircraft (remote sensing) observations are useful for a wide range of applications related to UNFCCC measuring, reporting and verification, including land cover/land-use change (activity data), aboveground biomass (AGB) assessments (i.e., carbon stocks and associated emission factors) and ‘stock change’ monitoring.

Challenges of measuring & monitoring with remote sensing
- Forest inventory and remote sensing methods of carbon stock assessment each have limitations but, importantly, are not at odds with one another. Rather, each approach can inform the other to generate better assessments than either approach alone.
- Remote sensing and classification of land cover modification by humans (activity data) is fraught with issues of classification ambiguities and the inability of satellites to detect land cover changes that can be attributed to human modification.
- Aircraft light detection and ranging (LiDAR) acquisitions are useful for capturing spatial variability across the landscape and for scaling between in situ (e.g., plot scale) measurements and satellite observations. They are very appropriate to project-scale activities and can be used to develop AGB estimates over regional to national scales when coupled with land cover/land-use classifications, but aircraft acquisitions are difficult to acquire systematically and repeatedly over very large areas, particularly if reporting requirements are frequent (e.g., biannually).

Opportunities for improved measuring & monitoring
- In situ measurements of carbon stocks can be used to calibrate and validate remote sensing observations, and inventory sampling approaches can be optimized by satellite maps (e.g., for stratifying sampling designs).
- LiDAR measurements are revolutionizing AGB estimates and thus not only stock change assessments but also estimates of other carbon pools (i.e., belowground biomass, litter, dead wood and organic soil), since the latter are typically derived, in part, from AGB.
- Radar, InSAR and optical imagery can be used synergistically with LiDAR to extend LiDAR ‘shots to wall-to-wall’ maps by developing statistical or physically based models to spatially extend local measurements. This approach is particularly powerful when the LiDAR measurements are calibrated with in situ observations.
- AGB maps derived from a combination of remote sensing and in situ data capture spatial variability across the landscape and thereby not only provide valuable estimates of carbon stocks but also the context for identifying and prioritizing areas with co-benefits (e.g., for biodiversity conservation) and setting baselines.
- Satellites are being developed specifically to improve estimates of AGB across the globe at sufficient spatial resolution to inform UNFCCC policies but also to reduce uncertainty in regional to global scale carbon budgets.

Bibliography
Papers of special note have been highlighted as:
- of interest
- of considerable interest
6 Maniatis D, Mollicone D. Options for sampling and stratification for national forest inventories to implement REDD+ under the UNFCCC. Carib. Bal. Manage. 5(9) (2010).
9 Outlines a plan for coordinating measurements and establishing an international network of data providers, systematic standards and protocols.
11 Provides an overview of various efforts to monitor tropical deforestation using satellite imagery.
15 Reports on an approach for capturing landscape scale variability in carbon stocks via coordinated satellite image classification, aircraft light detection and ranging (LiDAR) acquisitions and field measurements.
Overview paper describing the applicability of LiDAR applications in ecosystem science.


Uses LiDAR observations to create the first aboveground stock change estimate in a tropical forest ecosystem.


Website
101 Food and Agriculture Organization of the United Nations.
www.fao.org/forestry