Chapter 2. **Review of validation standards of Earth Observation derived biophysical products**

M Soto-Berelov *[1,2], S Jones [1, 2], E Farmer[1], W Woodgate[1,2]

[1] Southern Earth Observatory, School of Mathematical and Geospatial Sciences, RMIT University, Melbourne, Australia
[2] Cooperative Research Centre for Spatial Information, Victoria, Australia

*Corresponding author:
Mariela.soto-berelov@rmit.edu.au

In *AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data products*.

Abstract

In the context of remote sensing, validation refers to the process of assessing the uncertainty of higher level, satellite sensor derived products by analytical comparison to reference data, which is presumed to represent the true value of an attribute. Naturally, validation is an essential component of any earth observation program, since it enables the independent verification of the physical measurements obtained by a sensor as well as any derived products. After presenting some relevant definitions, this chapter draws on international and national validation campaigns to summarize some of the major components involved when using ground-reference data to validate biophysical products derived through Earth Observation (EO) data. These include site selection, site extent, and sampling design. The process of up-scaling, which enables the validation of coarse resolution products via the comparison of measurements made at various scales (i.e., ground-based, intermediate-airborne) is also reviewed. The chapter concludes with a brief section on alternative validation methods.

Key Points

- The Committee of Earth Observing Satellites has identified four stages of validation, each of which is progressively more comprehensive.
- Satellite derived products can be validated directly using an independent data source that is representative of the target values or indirectly through product inter-comparison and/or by collecting measurements across various scales and upscaling.
- Sites chosen for validation should meet certain criteria, including the following: be accessible to researchers; encompass existing facilities such as flux towers which collect measurements of biophysical variables over extended periods of time; have long-term commitment to scientific studies; represent significant areas of homogenous or uniformly mixed land cover.
- The site extent of a validation site must be large enough to represent the pixel size of the sensor being validated.
- The sampling design implemented for ground-based measurements is driven by two main factors: (a) the footprint of the field measurements and (b) up-scaling process used to integrate the field measurements and high resolution imagery
- Field activities should be carried out within a week of satellite/airborne acquisition to prevent significant changes in vegetation. However, the rate at which the state of the vegetation evolves varies for different ecosystems and is also influenced by its successional state.
- When devising a sampling design, many projects choose a sampling scheme based on elementary and secondary sampling units. Elementary sampling units (ESU) aim to capture the variability of the product being validated across the study site (this can be determined from a current land cover or floristic map, surface reflectance as characterized by recently acquired satellite imagery). Secondary sampling units (SSU) are distributed across the ESU and represent the specific locations where measurements are recorded. Different sampling designs can be implemented within SSU including fixed pattern, transect, randomized designs.
- Up-scaling is generally achieved via the integration of field measurements and a high-resolution image, which results in the production of a high resolution map of the parameter measured in the field.

2.1 Introduction

In 1984, the Committee on Earth Observing Satellites or CEOS was established following a recommendation by the Economic Summit of Industrialized Nations Working Group on Growth, Technology, and Employment's Panel of Experts on Satellite Remote Sensing (http://www.ceos.org), to coordinate space-borne observations across the planet that help address current and critical scientific research questions. CEOS endeavors to optimize the benefits of space-borne Earth Observation (EO) by planning missions through the collaborative participation of its members, which include space agencies and both national and international EO organizations. In addition, CEOS is directly involved in planning and developing accessible and compatible data products, formats, services, applications and policies (CEOS WGCV Work Plan 2011-2016, 2014) that relate to EO data and missions.

To be able to quantify data derived from EO missions and compare sensors and products and ultimately use these to tackle pressing scientific questions, CEOS established the Working Group on Calibration and Validation (WGCV) in 1984. The WGCV undertakes and promotes activities to coordinate and advance the calibration and validation of EO missions and data (Dowman 2004), so they can be of use across wide international user communities. When validating moderate resolution global products created from EO data such as MODIS, CEOS has identified four stages of validation (Table 2.1), each of which is progressively more comprehensive.

Table 2.1 CEOS Validation hierarchy (WWW2).

Stage	Description
Stage 1 Validation	Product accuracy has been estimated using a small number (typically < 30) of independent measurements obtained from selected locations and time periods and ground-truth/field program effort.
Stage 2 Validation	Product accuracy has been assessed over a widely distributed set of locations and time periods via several ground-truth and validation efforts. The spatial and temporal consistency of the product has been evaluated over globally representative locations and time periods. Results are published in peer-reviewed literature.
Stage 3 Validation	Product accuracy has been assessed over a globally distributed set of locations and time periods via several ground-truth and validation efforts. Product uncertainties have been well-established via independent measurements made in a systematic and statistically robust way that represents global conditions. Results are published in peer-reviewed literature.
Stage 4 Validation	Validation results for Stage 3 are systematically updated when new product versions are released and as the time-series expands.

The WGCV supports six subgroups. Each of these focuses on different technical areas (Table 2.2): land product validation; atmospheric composition; Synthetic Aperture Radar (SAR); microwave sensors; terrain mapping; infrared and visible optical sensors.

Table 2.2 Mission view of CEOS calibration and validation subgroups (CEOS WGCV five year working plan, 2012).

Subgroup	Mission
Land product validation	Foster quantitative validation of higher-level global land products derived from remote sensing data and report results so they are relevant to users.
Atmospheric composition	Ensure accurate and traceable calibration of remotely-sensed atmospheric composition radiance data and validation of higher level products, for application to atmospheric composition, land, ocean, and climate research.
Synthetic aperture radar	Foster high-quality synthetic aperture radar data from airborne and spaceborne systems through precision calibration in radiometry, phase and geometry, and validation of higher level products.
Microwave sensors	Foster high quality calibration and validation of microwave sensors for remote sensing purposes. These include both active and passive types, airborne and spaceborne sensors.
Terrain mapping	Ensure that characteristics of digital terrain models produced from Earth Observation sensors at global and regional scale are well understood and that products are validated and used for appropriate applications.
Infrared and visible optical sensors	Ensure high quality calibration and validation of infrared and visible optical data from Earth Observation satellites and validation of higher-level products.

Given this Handbook recommends guidelines for the validation of terrestrial biophysical products, the Land Product Validation (LPV) subgroup is of particular relevance. The LPV is also subdivided into focus areas that represent a key terrestrial Essential Climate Variables (ECV): land cover; fire; biophysical Leaf Area Index (LAI)[1]; Fraction Absorbed Photosynthetic Active Radiation (f_{APAR})[2]; surface radiation; land surface temperature; soil moisture; land surface phenology; and snow cover.

This chapter focuses on the field survey techniques utilized to collect validation data and only briefly considers the up-scaling of these recorded measurements. International validation campaigns from major earth observing programs such as MODLAND (MODIS land discipline team) and ESA VALERI (Validation of Land European Remote Sensing instruments, Baret et al., 2006) are used as examples, given their focus on validating medium resolution satellite products (i.e., MODIS, MERIS) related to land cover and vegetation (e.g., LAI, Foliage Projective Cover or FPC, Fractional Vegetation Cover or FVC).

2.1.1 Validation in Australia

Earth Observation data are critically important to a number of Australian research, environmental, and government monitoring programs. The reliability and use of such data depend on the extent to which such data have been calibrated and validated. International calibration and validation (Cal/Val) programs are biased towards northern hemisphere vegetated ecosystems, leaving many of Australia's unique terrestrial ecosystems under-represented. The Australian Academy of Technological Sciences and Engineering (AAS/AATSE) review of EO in Australia recognizes that Cal/Val of EOS data for the Australian region is a fundamentally important scientific activity. Accordingly, there is a need for EO data to be calibrated and

---

[1] Leaf area index or LAI is typically defined as the total one-sided area of leaf tissues per unit of ground surface area (Bréda 2003).

[2] f_{APAR} is defined as the fraction of photosynthetically active radiation (PAR) in the 400-700 nm wavelength range, that is absorbed by a canopy.

validated against high quality surface-based measurements across the continent following specific internationally agreed scientific criteria (AAS 2009).

Previous Australian involvement in international Cal/Val activities has been valuable; it has allowed Australian scientists to join international EOS science teams and has provided early access to important satellite data streams. Although there are major land cover monitoring exercises such as Australia's National Carbon Accounting System (NCAS) and the Queensland government State Land and Tree Survey or SLATS (Kuhnell et al., 1998) that have dedicated validation components, the national coordination and funding of Cal/Val activities has been limited and ad hoc in the past (AAS 2009).

During recent years, Cal/Val activities in Australia have been coordinated by the AusCover facility within TERN. AusCover is responsible for providing a new nationally consistent approach for collecting, validating and distributing biophysical products related to land cover and land surface (Figure 2.1) derived from time-series remote sensing systems. These products can then be used to support ecosystem research and resource management within Australia.

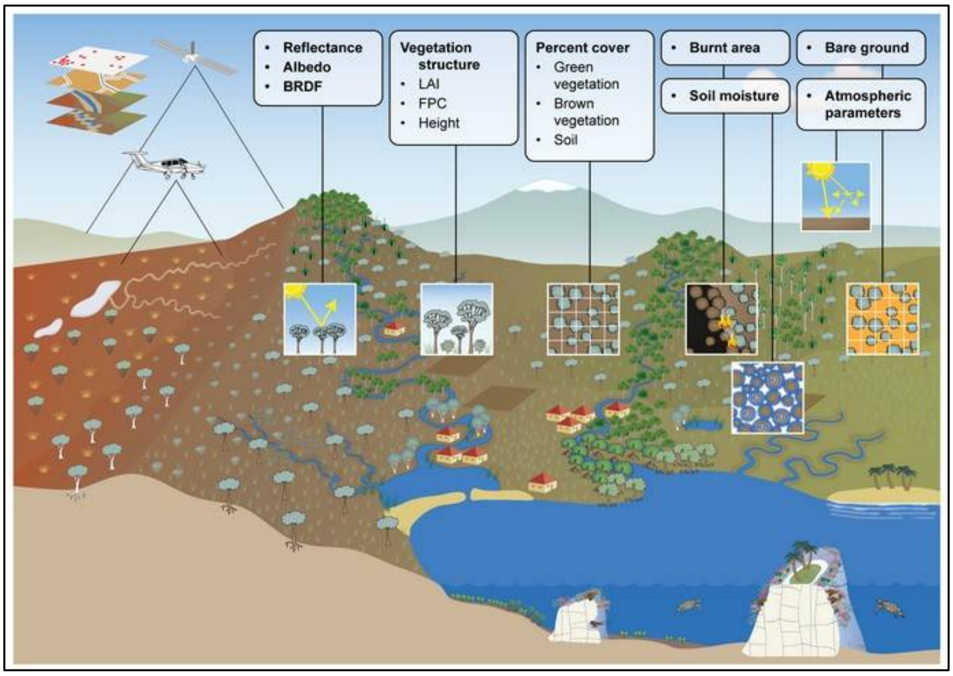

Figure 2.1 Representation of biophysical products provided by AusCover

AusCover has set up a national calibration and validation program to provide for the calibration and validation of biophysical products. In this context, AusCover validation activities aim to utilize independent field data, aerial and satellite data to assess the quality of a range of terrestrial land surface products (Figure 2.1). This assessment will contribute to Stage 4 validation (CEOS WGCV), the highest of the CEOS defined hierarchical validation levels (Table 2.1). At this level, validation aims to comprehensively establish product uncertainties via the utilization of independent measurements which are made in a systematic and statistically robust way and which are representative of global conditions.

Validation activities draw extensively on Cal/Val knowledge from international groups and campaigns like the CEOS WGCV, EOS-MODIS Bigfoot CAL/VAL, ESA VALERI and the National Ecological Observatory Network (NEON). In addition to the international expertise, AusCover also incorporates local knowledge from existing projects with dedicated validation schemes (e.g., NCAS, SLATS). Several examples of AusCover validation campaigns throughout TERN's National Scientific Reference Site Network (NSRSN) are given in Chapter 17. In Chapter 18, a nationally coordinated effort for the validation of fractional cover that is led by the Australian Bureau of Agricultural and Resource Economics and Sciences or ABARES is presented.

2.1.2 Terminology

In the context of remote sensing, validation refers to a process of assessing the uncertainty of higher level, satellite sensor derived products by analytical comparison to reference data, which is presumed to represent the target or true value of an attribute. To achieve this, conventional, ground-based observations are required using calibrated and traceable field instrumentation and associated methods. This allows for the verification and improvement of the algorithm/s used to derive the product. In a similar way, the CEOS WGCV defines validation as the process of assessing the uncertainty contained within satellite derived products via an analytical comparison to reference data (http://lpvs.gsfc.nasa.gov).

When validating a product, the accuracy or uncertainty contained within satellite derived products (e.g., land cover or Leaf Area Index) can be assessed directly or indirectly. Direct validation implies using an independent data source that is representative of the target values or surface conditions (Justice et al., 2000). This allows for an 'absolute' quantification of uncertainties. Unfortunately, direct validation is often limited by the number and quality of available reference data, thus limiting the spatial coverage. To counter this, products can be inter-compared (indirect validation) to provide an indication of gross differences and possible insights into the reasons for the differences (Justice et al., 2000). Such validation procedures consider (a) the internal spatial/temporal consistency of a data product; and (b) the consistency of a given data product relative to existing data products at a comparable spatial scale (i.e. inter-comparison). Although this has the potential to provide a more extensive evaluation of consistencies/differences between products, it lacks a link to quantitative reference data (direct validation).

In addition, products can also be validated indirectly through a two-stage process that involves the collection of measurements across various scales. At a large scale, ground observations are collected across an area that is representative of the resolution of the product that is being validated. The ground measurements can then be up-scaled to an intermediate scale using high resolution imagery (Morisette et al., 2006) and then compared to the product of interest.

2.2 Validation site requirements

The Moderate Resolution Imaging Spectroradiometer (MODIS) Land Discipline team (MODLAND), which leads validation efforts for MODIS derived biophysical products, has an established network of sites used for validation activities that is globally representative. In other words, sufficient sites were included to be representative of a given biome/ecosystem. Such representativeness was achieved by considering the distribution of sites both within the physical and meteorological space (Morisette et al., 2002). Despite a need for a globally representative framework, MODLAND recognized that given limited resources for data collection and analysis, the project should leverage on existing resources. This is achieved via the utilization of, and partnerships with, existing (a) field programs (such as Long-term Ecological Research sites or LTER); (b) science data networks (i.e., fluxnet); and (c) national and international research efforts (i.e., Morisette et al., 2002).

In establishing validation sites MODLAND established a series of criteria that define the optimum site location for satellite product validation (Morisette et al., 2002). According to these criteria, a validation site should:

- be accessible to researchers;
- encompass existing facilities such as flux towers, which collect measurements of biophysical variables over extended periods of time;

- have a long history and long-term commitment to scientific studies;
- represent significant areas of homogenous or uniformly mixed land cover;
- be representative of extensive biomes globally;
- be complementary to existing validation sites.

To validate products derived from medium resolution satellitesVALERI provides high spatial resolution maps of biophysical variables (e.g., LAI, fAPAR, fCover) that are estimated from ground measurements and high spatial resolution images like SPOT or Landsat ETM+. As part of their methodological framework, they rely on a network of sites distributed throughout the globe. These sites also need to be relatively homogenous (Baret et al 2006) within an area that is large enough to represent the spatial resolution of the sensor (at least 3km x 3km). In other words, variation in the biophysical variable of interest (and associated radiometric values) should be minimal across the study area extent (as you move from one area that represents $1km^2$ to another within the $3x3km^2$ site). Sites should also be representative of different biomes (dependent on available local support for field activities). Ideally, sites should also have relatively small topographic variation in order to simplify the interpretation of both the ground measurements and acquired satellite imagery (Baret et al., 2006).

In Australia, AusCover Cal/Val activities encompass an extensive large area validation campaign (>1000 sites) that takes advantage of sites surveyed by other facilities within TERN. An example is AusPlots Rangelands[3] which has established a network of permanent plots across rangeland areas throughout the Australian continent (spanning across 52 bioregions, Thackway and Cresswell 1995). Fractional cover and LAI measurements (along with other metrics) are recorded in these plots using the SLATS transect sampling method (discussed below). In addition, AusCover makes use of a series of sites that are intensively characterized (also referred to as super sites) and are suitable for multi-instrumental land product validation and algorithm development. Super sites are located across significant biomes within Australia and include representative areas of sclerophyll forests, savanna woodlands, grasslands and tropical forests (Figure 2.2).

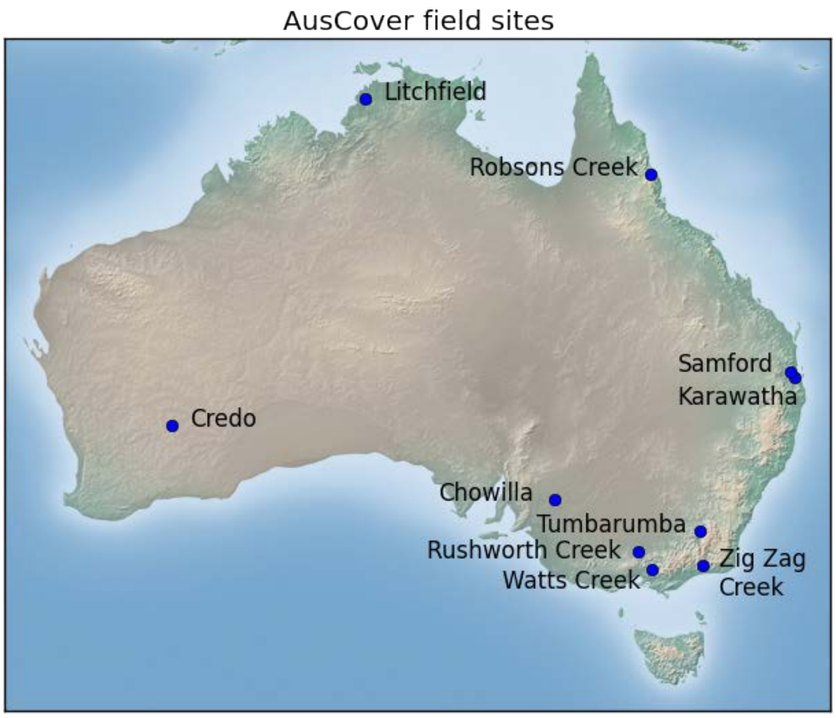

Figure 2.2.2 Network of supersites across Australia

[3] AusPlots rangelands is a sub-facility within the Terrestrial Ecosystem Research Network's Multi-scale Plot Network (MSPN) facility (http://www.tern.org.au/AusPlots-Rangelands-pg17871.html)

Some of the key criteria that need to be met for these sites to be chosen include:

- being representative of an important land cover and Australian biome;
- being spatially homogenous over a 5km x 5km footprint area so they can be scaled-up to validate large area remotely sensed products
- being easily accessible;
- wherever possible, incorporate existing research facilities (e.g., flux towers).

2.3 Site extent

Within current validation projects the site extent, or the area over which field measurement are collected, varies. Generally speaking, the extent of the site must ensure the representation of the pixel size of the sensor. The smallest possible site extent is the minimum area compatible with the spatial resolution of the sensor to be validated, typically 1 km^2 within current LAI products (Morisette et al., 2006). However, multiple authors conclude that a 1 km^2 area extent is too small given issues associated with the point spread function and geo-locational uncertainties of the sensors (Morisette et al., 2006). These issues have been minimized, in multiple studies, via the (a) positioning of sample sites within homogenous areas; and (b) definition of larger, typically 3km x 3km and 5km x 5km site extents.

When choosing the site extent, another important consideration is the available resources required for field work. Ideally, a site should be surveyed within a week of satellite/airborne image acquisition in order to prevent the significant evolution of the vegetation from the date of data capture (Baret et al., 2006) or occurrence of destructive events (e.g., fire). Nevertheless, the rate at which the state of the vegetation evolves varies for different ecosystems and is also influenced by its successional state.

In Australia, large area sites that are used for calibration and validation activities are 5km x 5km in extent. They are representative of different biomes and Australian forest ecosystems (Figure 2.2). Across these sites, airborne full waveform LiDAR and imaging spectroscopy data are collected synchronously with ground-based data using a variety of instruments that measure biophysical products like LAI, Canopy Cover (CC), FPC and allied metrics, and FPAR (see Chapter 18 for more information on several AusCover validation campaigns). An example schematic of a validation site and data collected during a validation campaign is shown in Figure 2.3.

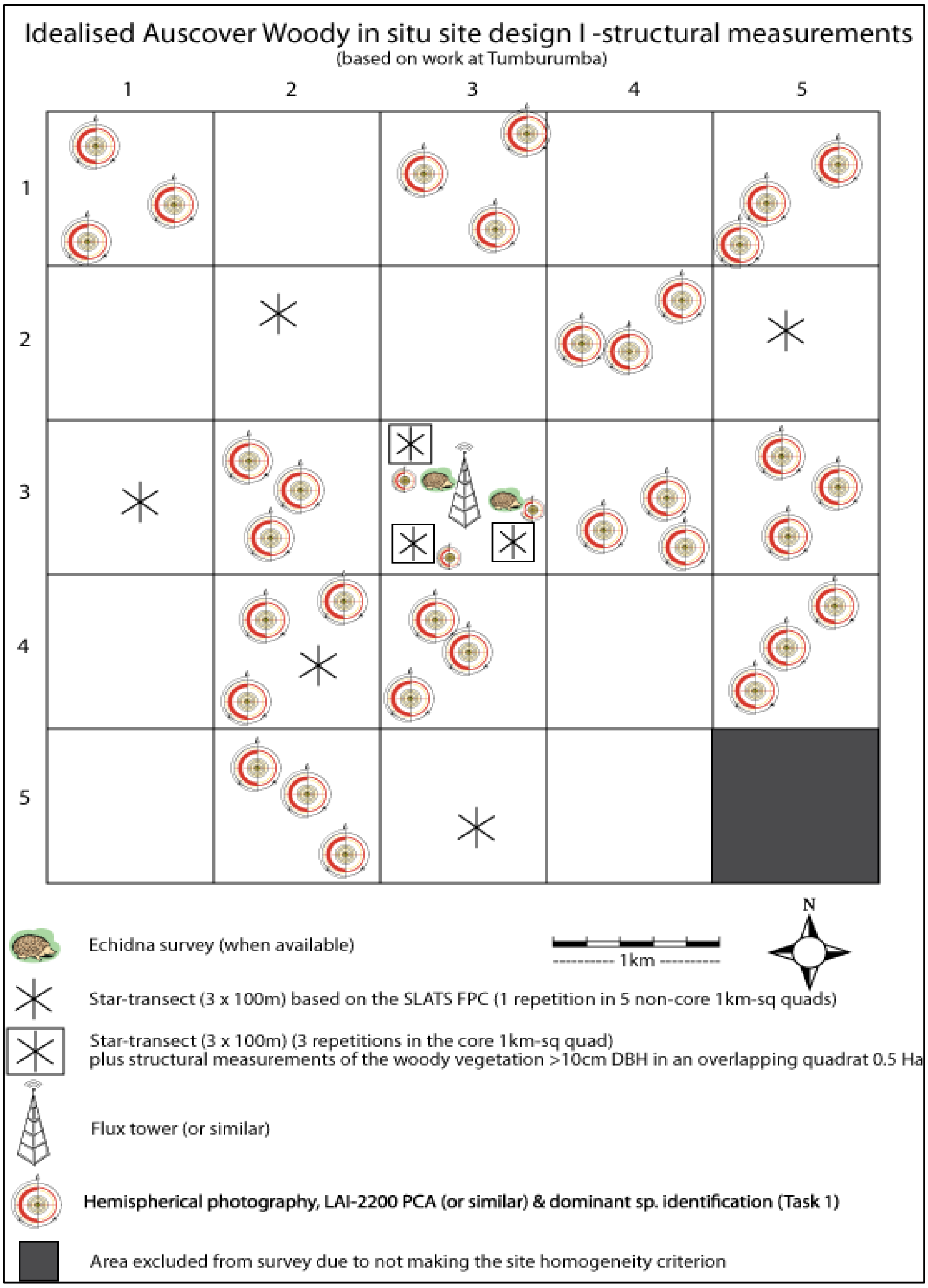

Figure 2.3 Schematic diagram of an AusCover large area validation site.

2.4 Sampling design

The sampling design implemented for ground-based measurements is driven by two main factors: (a) the footprint of the field measurements and (b) up-scaling process used to integrate the field measurements and high resolution imagery (see Table 6.3 for examples on sampling designs applied for LAI product validation). Multiple projects choose a multi-scale, two-tier sampling scheme based on elementary and secondary sampling units (e.g., Baret et al., 2006, Hufkens et al., 2008).

Elementary sampling units (ESU), also defined as primary sampling units, aim to capture the variability of the product being validated across the study site. The number and distribution of ESU across the study site varies between projects as a consequence of several factors including the site area, ESU extent, and site variability (Morisette et al., 2006). Site variability, within the ESU, can be defined according to a current land cover map, floristically, or using variability in the land surface reflectance as characterized by recently acquired satellite imagery.

Secondary sampling units (SSU) are distributed across the ESU and represent the specific locations where measurements are recorded. The distribution of these second-stage sampling units varies between projects and as a consequence of the (a) footprint of the product measurement device and (b) the land cover (canopy type) being studied. Different sampling designs can be implemented within the ESU (Morisette et al., 2006) such as those shown in Figure 2.4 (e.g., fixed pattern, Figure 2.4a; transect, Figure 2.4b; randomized design, Figure 2.4c).

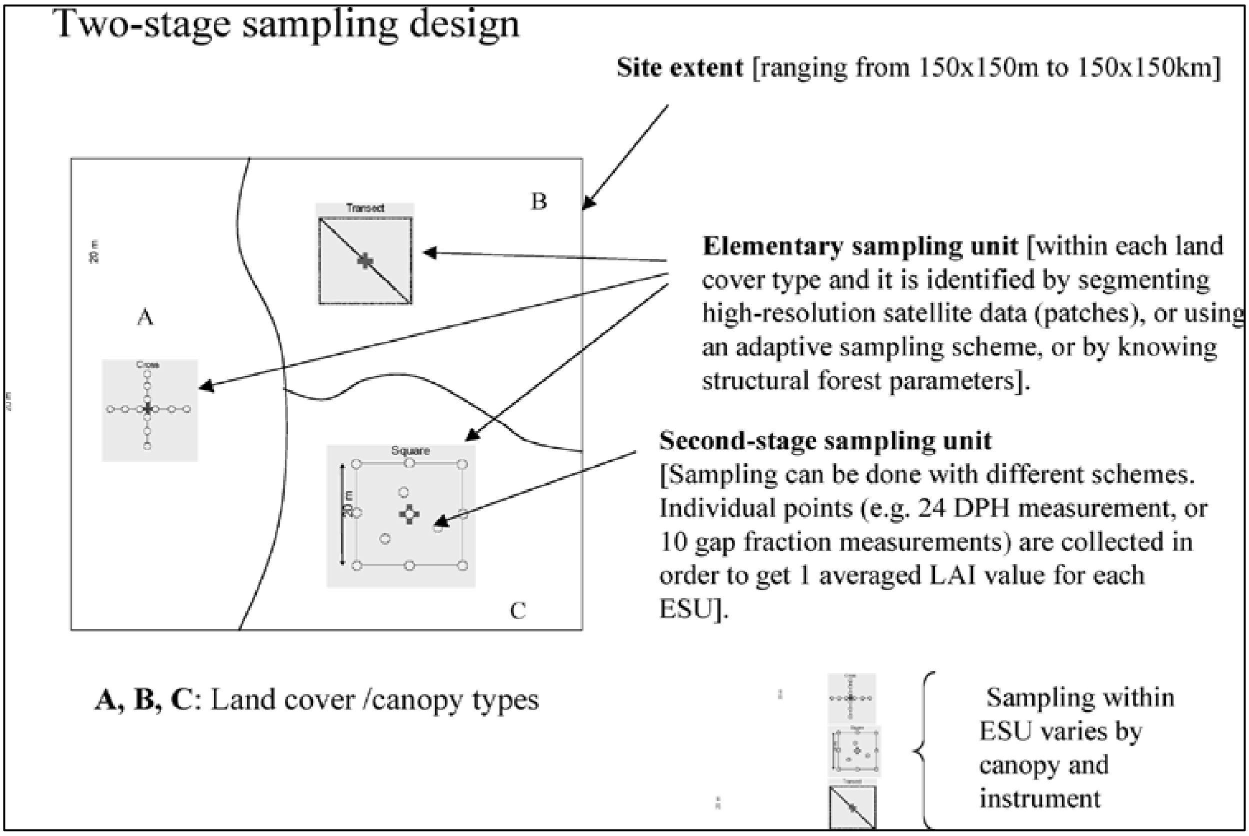

Figure 2.4 Commonly utilized two-stage sampling designs (Morisette et al., 2006) consisting of an Elementary sampling unit (ESU) and Secondary sampling unit (SSU).

Differences obtained when using different ESU have been investigated by some researchers. Garrigues et al (2002) found the fixed pattern (cross) sample design (Figure 2.4A) and randomized design (Figure 2.4C) to be equivalent in terms of the spatial variation sampled. In the case of LAI measurements, a transect design was recommended (a) with a TRAC device (Morisette et al., 2006) or (b) within land covers characterized by sparse or locally discontinuous vegetation (Baret et al., 2006). The second approach is particularly evident in the VALERI program in which the transect sample design was associated with destructive (as opposed to relying on indirect methods through DHP or instruments like LAI-2200) measurements of LAI (See Chapter 6 for detailed information on LAI validation).

A GPS is usually needed to record the precise location of ESU (where all measurements are taken or at the centre of the ESU). The specification of these GPS measurements and associated positional error is a function of the project and of course the resolution of the pixel. For example, VALERI utilized non-differential GPS to locate the centre of the ESU with an error of 5 to 10 metres.

2.4.1 VALERI

For the validation of LAI products derived from medium resolution satellite sensors (Instantaneous Field of View or IFOV=250-300m) within the VALERI project, validation sites were defined to cover a 3km x 3km area of homogenous or constantly mixed land cover (Baret et al., 2006). Between 27 and 45 ESU were located (as a function of study site) within the 9 km^2 site. ESU were defined to be 20m x 20m in extent, given they are mainly using SPOT-HRV satellite images for upscaling (with a spatial resolution that ranges between 10 and 20 metres).

To guarantee a good distribution of ESU across the site, the 9 km^2 study area was subdivided into 1 km^2 tiles with three to five ESU contained in each tile. The distribution of ESU within the 1 km^2 tiles was a function of (a) land cover; (b) ESU variability; (c) access; and (d) existing ESU locations, that is, ESU were required to be well distributed within the individual 1 km^2 tile. The representativeness of this sampling design was ensured at each site via a comparison of the Normalized Difference Vegetation Index or NDVI distribution (extracted from the high resolution imagery) of the entire site to that of the sampled ESU (Baret et al 2006).

The spatial distribution of biophysical measurements, within each ESU of 20m x 20m, was primarily a function of the dominant vegetation and its canopy structure (Baret et al., 2006). If the vegetation was considered to be homogenous, estimates of LAI were made using gap fraction techniques arranged in a cross or square spatial sampling distribution (Figure 2.1*a* and *c*). Conversely, a transect sample placed diagonally across the ESU (Figure 2.1*b*) was implemented if vegetation in the ESU was considered to be heterogeneous (Baret et al., 2006).

2.4.2 BigFoot

To support the validation of land products derived from MODIS such as LAI, land cover, and Net Primary Production (NPP), NASA's Terrestrial Ecology Program developed the BigFoot project (https://daac.ornl.gov/BIGFOOT_VAL/bigfoot.shtml). Sites chosen for validation typically have a 5km x 5km extent and have a flux tower in the centre (Figure 2.5) which measures water and carbon fluxes over a 1km^2 footprint to characterize NPP (Cohen et al., 2006). To evaluate the inter-annual validity of MODIS products, measurements of ecosystem structure and function are collected throughout the year.

Seven of the nine BigFoot sites were characterized by a sample design that included approximately 100 ESU (also termed plots), each 25m x 25m where measurements were collected (Cohen et al., 2006). The extent of the ESU approximates a Landsat pixel, which is the high-resolution satellite image used to up-scale ground-based measurements. To ensure the adequate characterization of vegetation within the flux tower footprint, between 60 and 80 ESU were concentrated in the 1 km^2 cell surrounding the flux tower. Remaining ESU were located within the 5KM x 5 km site (Cohen et al., 2006). ESU across the greater site extent were apportioned between the basic land cover components to enable the validation of BigFoot surface products over the full site. Conversely, within the centre 1 km^2 area, the ESU were sampled using a systematic spatial-cluster design (Burrows et al., 2002).

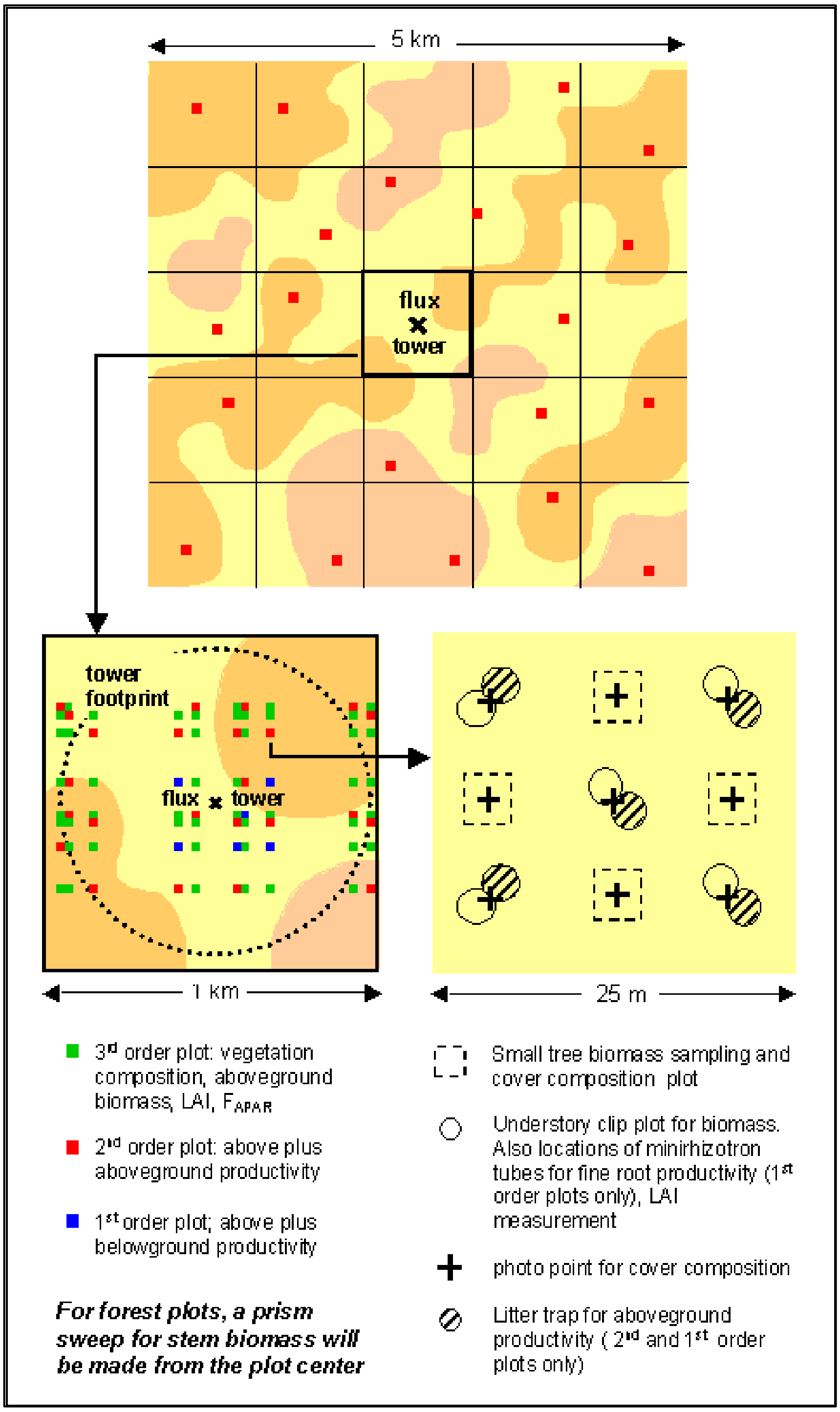

Figure 2.5 Overview of the BigFoot sample design (BigFoot Website: http://www.fsl.orst.edu/larse/bigfoot/ovr_dsgn.html).

It should be noted that a modified sampling methodology to that just outlined was developed for two of nine BigFoot sites (Kennedy et al., 2002). This modified sampling methodology was based on the use of 42 intensive ESU and 58 extensive ESU. This approach used Landsat ETM+ data to roughly characterize the range of conditions within the site in order to enable the efficient and effective allocation of samples. Within this sample design, the placement of samples was required to meet three objectives: (a) sufficiency (capture variability across the landscape); (b) efficiency (minimize field travel costs and expenses); and (c) independence of observation, therefore avoiding replication in the field data (Kennedy et al., 2002). This was achieved using a constrained stochastic sampling protocol for the placement of extensive samples in the greater site area (Kennedy et al., 2002).

As shown in Figure 2.5, the BigFoot ESU are characterized using a multi-tiered hierarchy; that is, plots were sampled at three levels of intensity. Measurements collected within each ESU were a function of their hierarchical classification. At the lowest hierarchical level (third order plots), measurements of vegetation composition, aboveground biomass, LAI and f_{APAR} were taken. These measurements were repeated at second order plots with the addition of above ground productivity. In the highest first order plots, all third order measurements are collected in addition to above and below ground productivity (Figure 2.5). All ESU in the greater 5km x 5km site footprint are characterized with second order measurements. However, the proportion of each ESU or plot type is a function of the site, as exemplified by Table 2.3.

Table **2.3** Example of hierarchy of ESU in two BigFoot sites (Campbell et al., 1999). NOBS (Northern Old Black Spruce) is a boreal forest site characterized by black spruce while KONZ (Konza Prairie) is a tall grass prairie site.

SITE	First Order	Second Order	Third Order	Total
NOBS	8	44	56	108
KONZ	6	38	56	100

Each ESU contains a series of sub-plots where the measurements described above are collected. Sub-plot placement was designed to ensure (a) the spatial stratification of measurements throughout the plot; (b) simple and convenient field deployment; and (c) minimal interference between the required measurements, for example, direct and indirect measurements of LAI (Campbell et al., 1999). The arrangement of sub-plots typically follows a regular pattern approximating the compass cardinals (Figure 2.6). However, sub-plot arrangement varies as a function of the vegetation type and site characteristics.

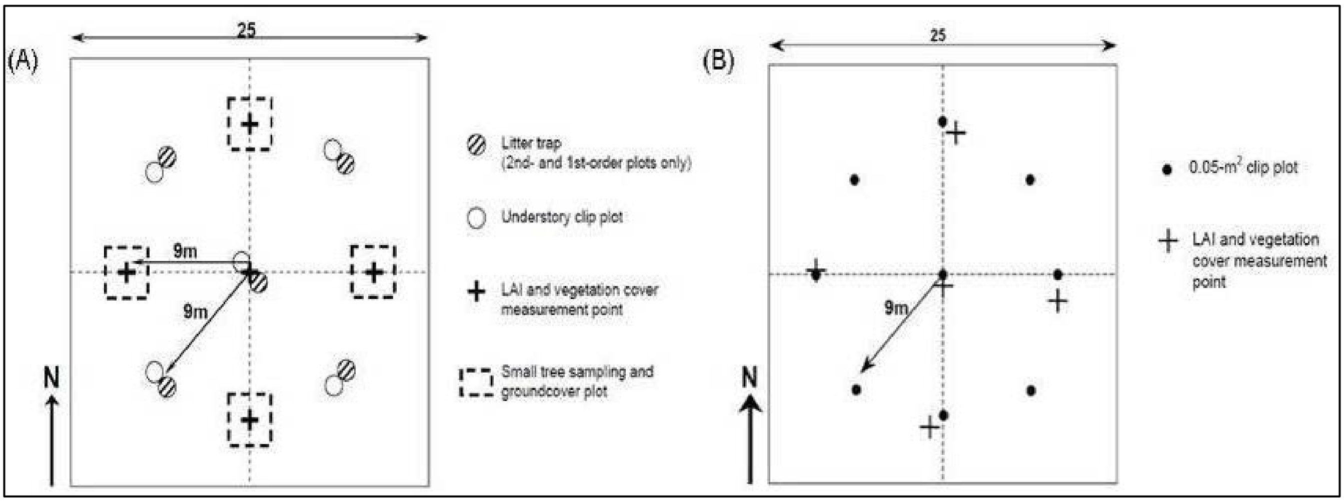

Figure 2.6 The arrangement of sub-plots and LAI measurements taken in the (a) NOBS and (b) KONZ BigFoot sites (Campbell et al., 1999).

2.4.3 SLATS

In Australia, a sampling strategy widely used to collect field data for calibrating and validating fractional cover products is SLATS, which also provides information on land clearing, tree growth and regrowth (Kuhnell et al., 1998; Muir et al., 2011). The SLATS sampling method has proved to be robust when using medium resolution products like Landsat for up-scaling in relatively open ecosystems across Australia. Woody vegetation mapping, within the SLATS project, is based on the automated and semi-automated classification of Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) satellite imagery (Kuhnell et al., 1998). According to the SLATS method, field sites are located in areas of uniform, mature vegetation communities (based on aerial photographs) and have a minimum area of approximately 100 x 100 metres (Armston et al., 2009). Three line segments, each 100 m in length, are orientated at 0°, 60° and 120° from magnetic north.

Vegetation characteristics recorded at SLATS sites include foliage projective cover (FPC) and stand basal area (SBA). Estimates of over-storey FPC are derived by averaging across three 100 metre point intercept transects at one metre spacings (Figure 2.7). At one metre intervals along each transect, overstorey (woody plants greater than two metres in height) and understorey (woody and herbaceous plants less than two metres in height) plants are recorded. Understorey herbaceous measurements are acquired using a laser pointer at zenith of zero (nadir) with intercepts classified as (a) green leaf; (b) dead leaf; (c) bare; (d) rock; (e) cryptogam; or (f) litter by the observer. Over and understorey woody measurements are done via a vertical tube method with intercepts classified as (a) green leaf; (b) dead leaf; (c) woody branch or stem; or (d) sky. Stand basal area measurements are collected at the centre point of the SLATS transect and at a 25 metre distance from the centre location along each of the line segments. Stand basal area is estimated, for each plot, as the average of seven optical wedge counts, with the transect representing the centre of a nominal one hectare plot (Figure 2.7). As described elsewhere in this handbook (e.g., Chapters 6, 12, 17), the SLATS transect can also be used to collect measurements of LAI and other metrics using a variety of instruments.

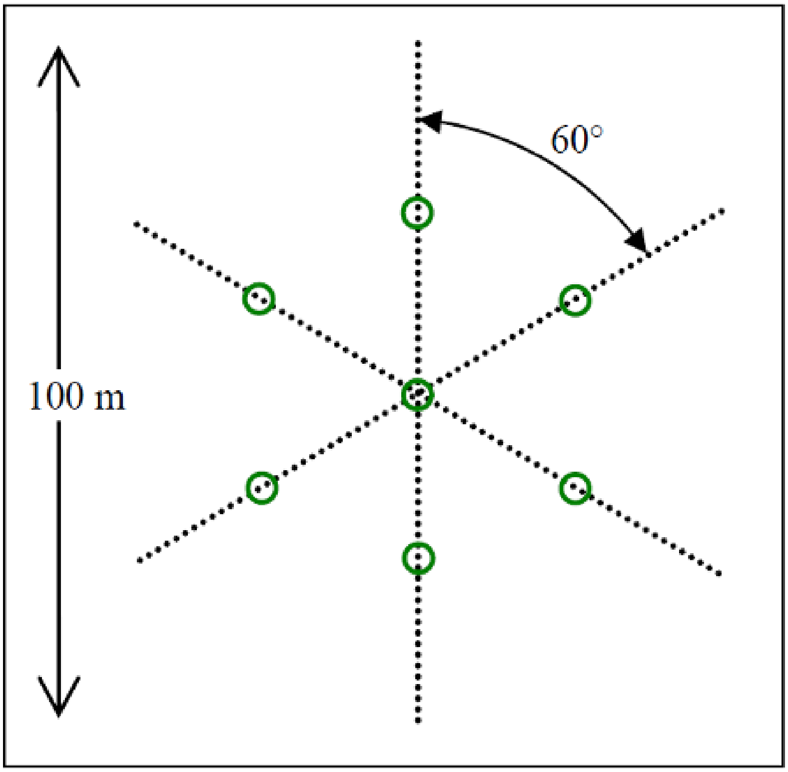

Figure 2.7 Schematic representation of the sampling design utilised in the SLATS survey (Armston et al., 2009). Green circles are located halfway between the centre location and end point (in other words, 25 metres from the central location) of each 100 m segment. Stand basal area measurements are collected in the green circles, where measurements of LAI and other metrics can also be collected.

Woodgate el al. (2012) recently compared plot scale LAI and FPC measurements obtained when using various sampling designs in a rainforest in Queensland (Figure 2.8). Three sampling designs were compared, namely SLATS, the VALERI cross, and a gridded one hectare plot sampled every 20 m. Their preliminary findings suggest that, in dense canopy forests, measurements obtained using various sampling designs are highly comparable and therefore the selection of the optimal sampling design should be driven by the resolution of the product that is to be validated. Nevertheless, additional factors should still be considered such as the type of forest, accessibility, and topography. For instance, when working in forests with dense canopies and understory vegetation, it may be extremely time consuming to conduct a SLATS transect and other sampling designs such as a simple transect or modification of one of the more widely used sampling designs may be more suitable (see Chapter 17, section 17.4, TERN AusCover field and airborn campaign in the wet tropics of Far North Queensland for an example).

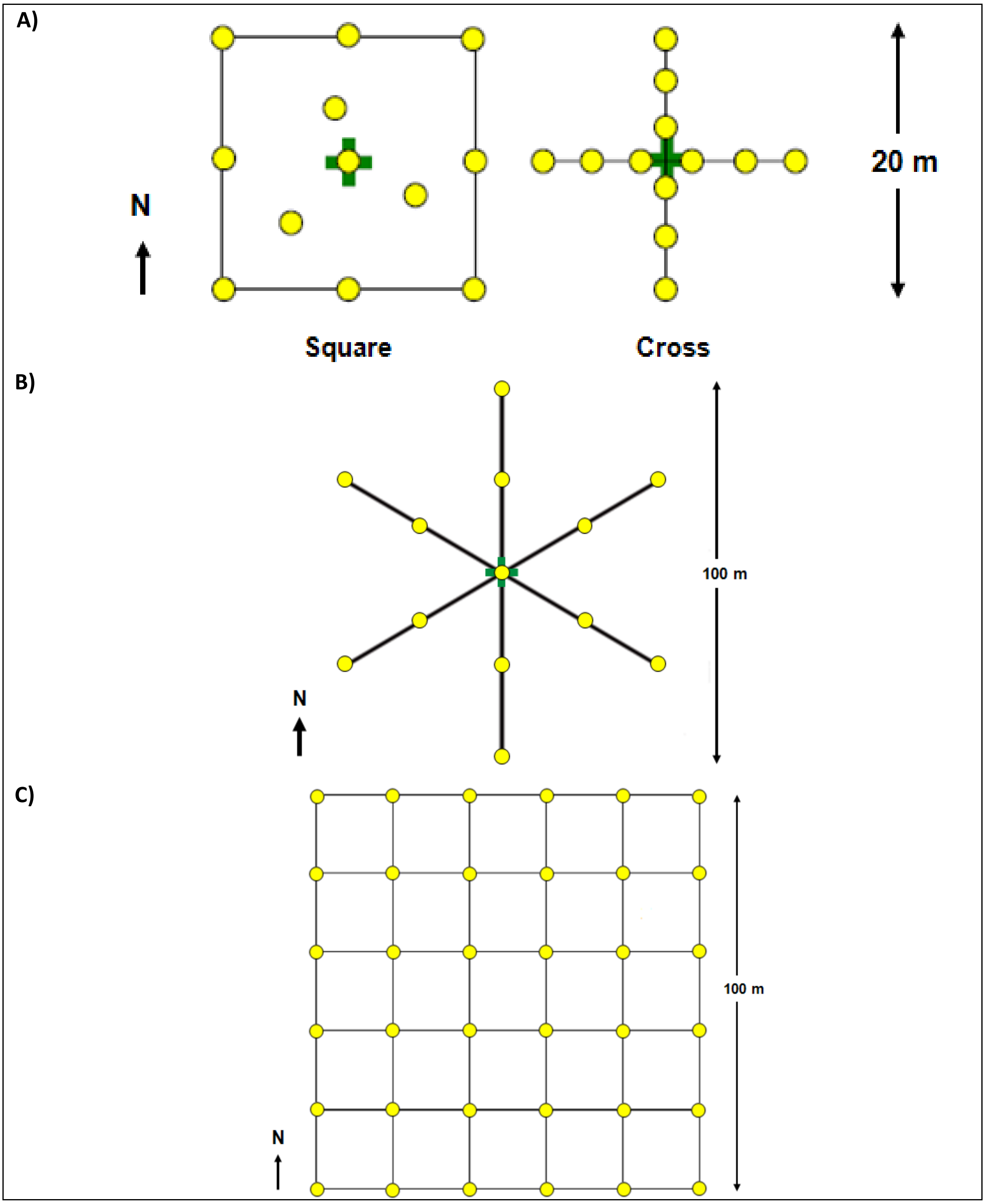

Figure 2.8 Three sampling designs investigated by Woodgate et al., (2012) where yellow dots indicate locations where LAI measurements were taken: A) VALERI cross, the green cross represents the centre of the plot which also has a GPS or known location associated with it. B) SLATS plot, C) Grid one hectare plot sampled every 20m.

2.5 Multi-stage sampling and upscaling

Validation procedures developed for moderate resolution satellite derived products emphasize the utilization of multi-scaled approaches which integrate ground-based, airborne, and high spatial resolution satellite data collected in tandem (Morisette et al., 2006). Ground-based (plot scale) data can be used to validate moderate spatial resolution remote sensing models by extrapolating the field measurements to a continuous spatial area that has a compatible scale with the spatial resolution of the remotely sensed observations (Baccini et al., 2007). It is in this context that ground-based measurements are collected in tandem with airborne or higher spatial resolution satellite imagery, which is used as a bridging data source between the field data and land product requiring validation. Chapter 17 presents a series of AusCover validation campaigns that involved in situ ground-data collection activities concurrent with the capture of high resolution hyperspectral and LIDAR data.

Up-scaling is generally achieved through the integration of field measurements and a high-resolution image, which results in the production of a high resolution map of the parameter measured in the field (Figure 2.9). A crucial consideration when designing the sampling framework is to embed the observations in a way that allows for them to be up-scaled, from point observations to landscapes to regions to continents. Validation of the moderate-resolution product is then achieved via comparison to this high resolution product (Morisette et al., 2006). Chapter 7 discusses the validation of an Australian national Fractional Cover product using MODIS and Landsat.

Advanced methodological techniques capable of supporting the up-scaling of ground-based measurements to continuous high resolution maps is an area of extensive research. In the case of LAI validation for example, projects utilize a range of (a) high resolution data sources; (b) transfer functions to integrate the ground and high resolution data sources; and (c) different procedures to validate the high resolution product. A review of such methods is beyond the scope of this document (for detailed reviews see, for example, Baccini et al., 2007; Baret et al., 2006; Gupta et al. 1998; Hay et al. 1997; Hay et al. 2001).

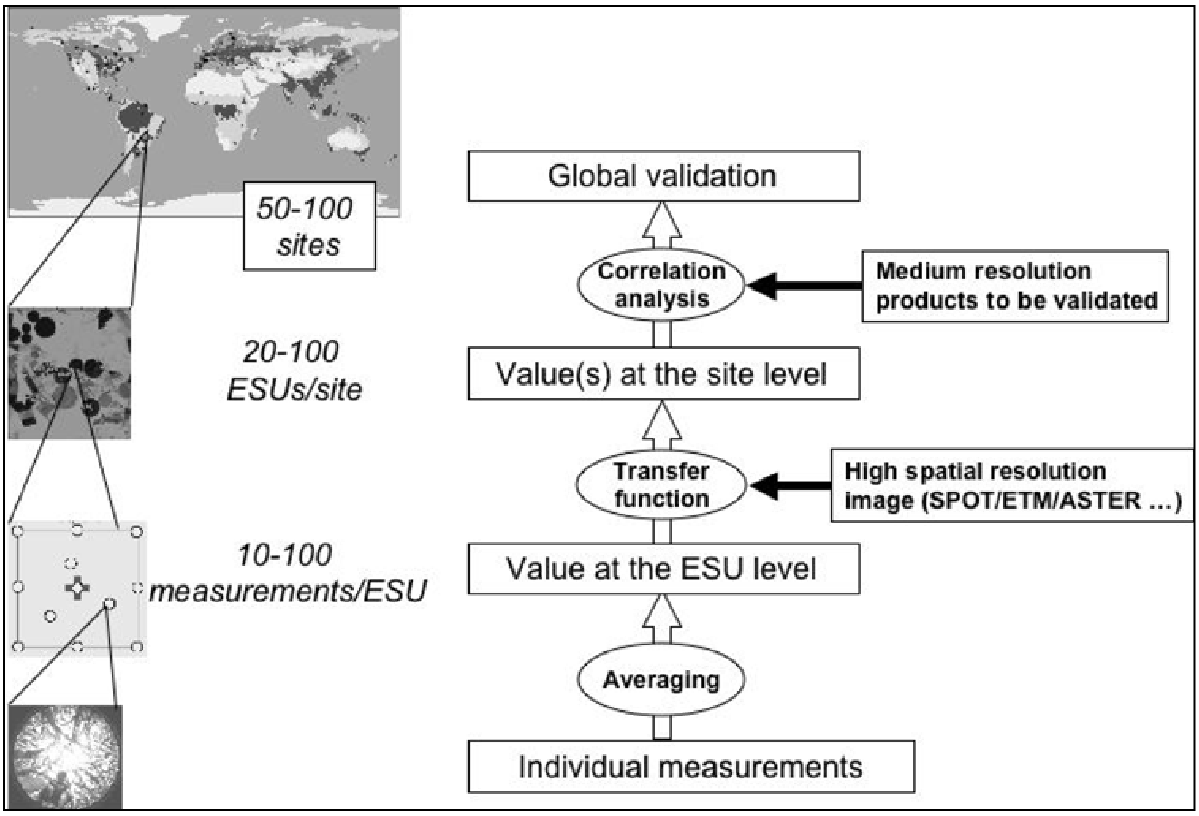

Figure 2.9 Validation and up-scaling procedures (from Morisette et al., 2006).

2.6 **Alternative validation approaches**

The validation of satellite derived products is necessary so that an estimate of the product's accuracy can be provided. Take the land cover product, for instance. Nowadays, multiple global and regional land cover products exist. Several authors have commented that the independent accuracy assessment of each of these products is inefficient, expensive and, due to the variety of validation procedures utilized, hinders the comparison of map accuracies (Stehman et al., 2010). Given these concerns there is an increasing move, within land cover mapping, towards a coordinated global land cover validation database (Stehman et al., 2010). Fundamental to this coordinated validation database is a rigorous probability sample of reference land cover data, which must (a) be compatible with all land cover class definitions; and (b) be based on a consistent response (sample) design protocol (Stehman et al., 2010).

The basic spatial unit of the proposed validation dataset is a 5km x 5km block (Stehman et al., 2010). It is proposed that reference land cover data be derived in each sample block from a high-resolution data source. It is this reference dataset that will form the basis of map comparison and therefore land cover map validation (Stehman et al., 2010). The sample design in which each of these blocks will be placed is required to: (a) represent a probability based sampling design; (b) adequately sample rare land cover classes; and (c) allow flexibility to easily augment the sample, via the sampling of particular regions or strata, to tailor the available samples to the assessment of a particular land cover product (Stehman et al., 2010).

To fulfill these criteria Stehman et al (2010) propose a stratified random sampling design. To avoid sampling bias towards a particular schema, the strata within this design would not be based on a single land cover representation but utilize instead a more generalized stratification based on Koppen climatic zones and population density (Stehman et al., 2010). The initial sample is based on 500 blocks allocated to the 21 strata to ensure that complex, potentially ambiguous classes receive a higher proportion of samples (Stehman et al., 2010). The augmentation of these original samples will be based on the same probability based sample design and original strata. However, it is expected that users could increase the sampling of strata, known to contain land cover types of interest, while still maintaining a probability based sampling approach (Stehman et al., 2010). This sample design is flexible and allows the addition of new samples targeted to the validation of a particular land cover product. Such characteristics are particularly relevant to the current review.

Another approach to validation is the large-scale transect method, which biases the sampling along an environmental gradient (e.g., elevation, temperature, precipitation). The advantage of this approach is that it encompasses a large range of environmental conditions but has the disadvantage of not being representative of common ecosystem "states". An example of large-scale transects are those established by the International Geosphere-Biosphere Prgram (IGBP), which extend for over 1,000 kilometres and span across different biomes (Koch et al., 1995). Figure 2.10a shows the distribution of IGBP terrestrial transects, scattered across four main regions (high and mid latitudes, semi-arid tropics, and humid/sub-humid tropics). Only one of these, the NATT, falls within Australia (Sea et al., 2011 collected ground observations along this transect to validate the MODIS MC4 and 5 LAI products). Nevertheless, the Australian Transect Network Subfacility within TERN's Multi-Scale Plot Network facility has established other transects across the Australian continent which also extend for hundreds of kilometers (Figure 2.10b).

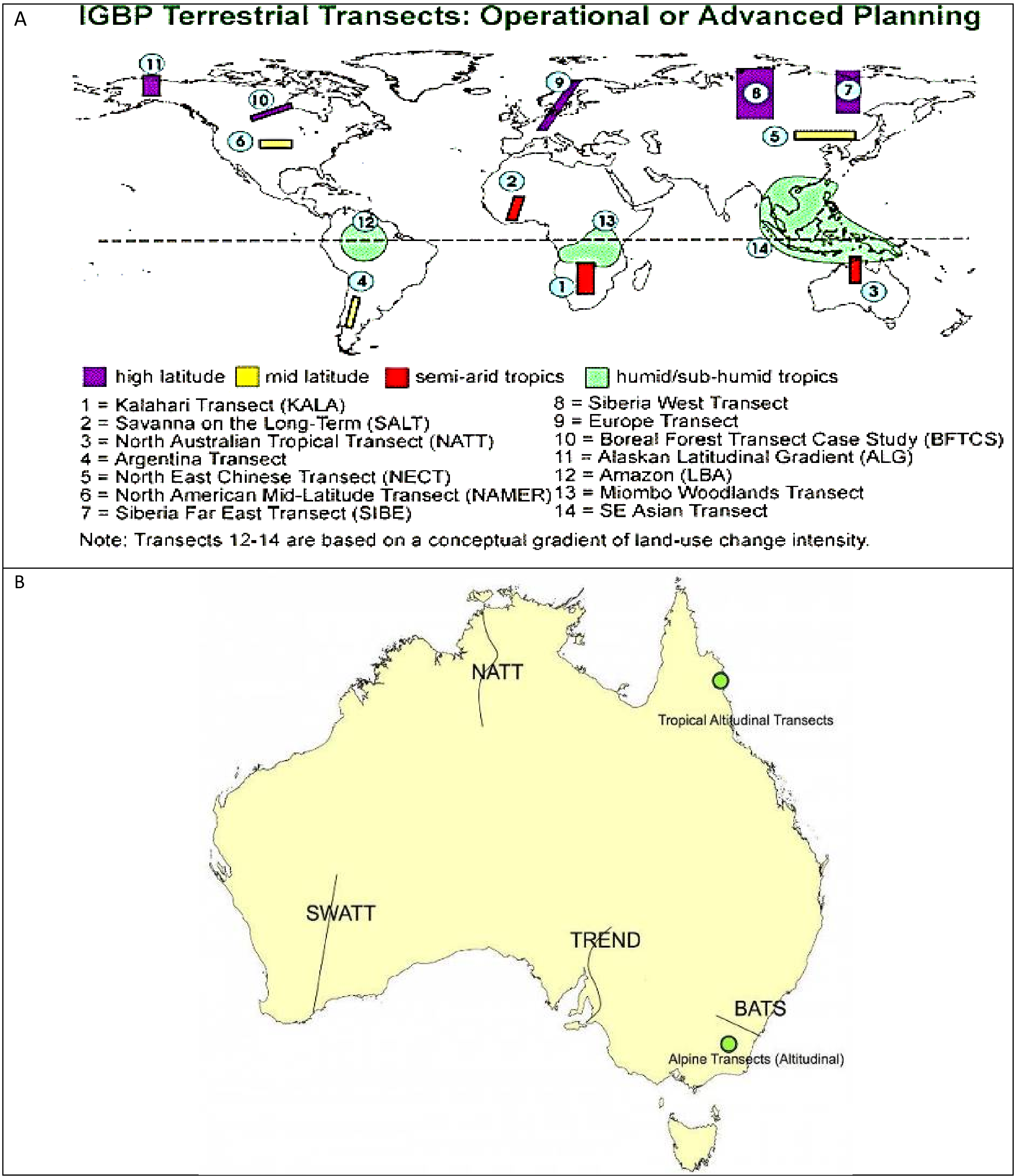

Figure 2.10 Examples of transects along environmental gradients. A) A set of IGBP Terrestrial transects. B) Four major transects that are part of the Australian Transect Network (http://www.tern.org.au/Australian-Transect-Network-pg22748.html).

2.7 Conclusion

There is a clear need for satellite remote sensing data to be validated to ensure the continued long-term provision of reliable datasets and products. Validation is a fundamentally important scientific activity. It needs to be an almost continuous component operating in tandem with EO campaigns that provides an independent check on the performance of space-based sensors and processing algorithms using high quality surface-based measurements and adhering to international guidelines and protocols. In Australia, this activity is being coordinated and implemented by AusCover, who are collecting data across multiple spatial scales (ground based, airborne, and satellite data).

The successful implementation of validation activities requires early and careful planning. The sampling framework needs to be practical and consider issues like site selection (potentially selection and establishment of networks of sites), size extent, sampling framework, coordination of sampling activities, and the development and deployment of required instrumentation. This chapter has briefly discussed some of these aspects by reviewing international validation campaigns as well as by drawing on national Australia validation campaign efforts. Other chapters in this handbook provide good practice guidelines when collecting other data in the field such as phonological measurements (Chapter 9), biomass (Chapter 12) and vegetation spectroscopy (Chapter 13). When embarking on a field campaign, it is important to consider all the attributes that will be measured in the field and logistics associated with acquiring these.

Not reviewed in this chapter but also of utmost importance to in situ data collection is quality assurance and data quality aspects. Data quality elements such as positional and attribute accuracy, logical consistency, and completeness need to be recorded for all in situ measurements using agreed upon protocols and standards (Chapter 3). In addition, Chapter 17 (TERN AusCover Field and Airborne Campaign at Robson Creek -Wet Tropics, Far North Queensland -Need to add chapter number later) presents several Cal/Val AusCover field/aerial acquisition campaigns as case studies, which demonstrate the logistics (e.g., number of personnel needed to record measurements of various metrics, instruments required, sampling strategy) required to undertake such work.

References

Baret, F., Morissette, J.T., Fernandes, R.A., Champeaux, J.-L., Myneni, R.B., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S., Nickeson, J.E. (2006). Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: proposition of the CEOS-BELMANIP. Geoscience and Remote Sensing, IEEE Transactions on , vol.44, no.7, doi: 10.1109/TGRS.2006.876030, available at http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1645280&isnumber=34482

Australian Academy of Science, Australian Academy of Technological Sciences and Engineering. (2009) An Australian Strategic Plan for Earth Observations from Space. Australian Academy of Science 2009 GPO Box 783, Canberra, ACT 2601, Australia. ISBN 085847 267 8, www.science.org.au/reports/index

Baccini, A., Friedl, M.A., Woodcock, C.E. & Zhu, z. (2007). Scaling Field Data to Calibrate and Validate Moderate Spatial Resolution Remote Sensing Models, Photogrammetric Engineering & Remote Sensing 73(8) p. 945 - 954.

Campbell, J., S. Burrows, S. Gower, and C. WB. (1999). Bigfoot: Characterizing land cover, LAI, and NPP at the Landscape Scale for EOS/MODIS Validation. Field Manual 2.1. Oak Ridge National Laboratory, Environmental Science Division.

Committee on Earth Observation Satellites Working Group on Calibration and Validation (CEOS WGCV) Five-Year Work Plan 2011-2016. (2014). Version 5.5.
http://ceos.org/document_management/Working_Groups/WGCV/WGCV_5-Year-Work-Plan-2011-2016_Feb2014.pdf.

Dowman, I. (2004) “Foreword,” in International Society for Photogrammetry and Remote Sensing (ISPRS) Book Series-Post-Launch Calibration of Satellite Sensors, S. A. Morain and A. M. Buedge, Eds. Leiden, The Netherlands: A. A. Balkema.

Garrigues, S., Allard, D., Weiss, M. & Baret, F. (2002). Comparing VALERI sampling schemes to better represent high spatial resolution satellite pixel from ground measurements: How to characterise an ESU, http://w3.avignon.inra.fr/valeri/methodology/samplingschemes.pdf (accessed August 2010).

Gupta, R. K., Prasad, S., & P.V., K. R. (1998). Evaluation of spatial upscaling algorithms for different land cover types. *Adv. Space Res*., 22(5), 625–628.

Hay, G. J., Niemann, K. 0., & Goodenough, D. G. (1997). Spatial Thresholds, Image-Objects, and Upscaling: A Multiscale Evaluation. *Remote Sensing of Environment*, 62, 1–19.

Hay, G. J., Marceau, D. J., & Dub, P. (2001). A multiscale framework for landscape analysis : Object-specific analysis and upscaling. *Landscape Ecology*, 16, 471–490.

Hufkens, K., Bogaert, J., Dong, Q. H., Lu, L., Huang, C. L., Ma, M. G., Che, T., et al., (2008). Impacts and uncertainties of upscaling of remote-sensing data validation for a semi-arid woodland. Journal of Arid Environments, 72(8), 1490–1505. doi:10.1016/j.jaridenv.2008.02.012.

Justice, C., Belward, A., Morisette, J., Leiws, O., Privette, J. & Baret, F. (2000). Developments in the ‘validation’ of satellite sensor products for the study of the land surface, International Journal of Remote Sensing, 21 (17) p. 3383 - 3390.

Kuhnell, C.A., Goulevitch, B.M., Danaher, T.J. & Harris, D.P. (1998). Mapping Woody Vegetation Cover over the State of Queensland using Landsat TM imagery, 9th Australian Remote Sensing and Photogrammetry Conference, Sydney, Australia.

Morisette, J.T., Privette, J.L. & justice, C.O. (2002). A framework for the validation of MODIS Land products, Remote Sensing of Environment, 83 p. 77 - 96.

Morisette, J.T.; Baret, F.; Privette, J.L.; Myneni, R.B.; Nickeson, J.E.; Garrigues, S.; Shabanov, N.V.; Weiss, M.; Fernandes, R.A.; Leblanc, S.G.; Kalacska, M.; Sanchez-Azofeifa, G.A.; Chubey, M.; Rivard, B.; Stenberg, P.; Rautiainen, M.; Voipio, P.; Manninen, T.; Pilant, A.N.; Lewis, T.E.; Iiames, J.S.; Colombo, R.; Meroni, M.; Busetto, L.; Cohen, W.B.; Turner, D.P.; Warner, E.D.; Petersen, G. W.; Seufert, G.; Cook, R. 2006. "Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup," Geoscience and Remote Sensing, IEEE Transactions on , vol.44, no.7, pp.1804,1817, available at http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1645281&isnumber=34482

Muir, J, M Schmidt, D Tindall, R Trevithick, P Scarth and J Stewart (2011). Field measurement of fractional ground cover: A technical handbook supporting ground cover monitoring for Australia. prepared by DERM for ABARES. Canberra.

Sea, W. B., Choler, P., Beringer, J., Weinmann, R. a., Hutley, L. B., & Leuning, R. (2011). Documenting improvement in leaf area index estimates from MODIS using hemispherical photos for Australian savannas. Agricultural and Forest Meteorology, 151(11), 1453–1461.

Stehman, S., Olofsson, P., Woodcock, C., Friedl, M., Sibley, A., Newell, J., Sulla-Menashe, D. & Herold, M. (2010). Designing a reference validation database for accuracy assessment of land cover, Accuracy 2010 Symposium, July 20-23, Leicester, UK.

Thackway R and Cresswell I (Eds) 1995, An interim Biogeographic Regionalisation for Australia; A Framework for Establishing the National System of Reserves, version 4.0, Australian Nature Conservation Agency, Canberra.

Woodgate, W, M Soto-Berelov, L Suarez, S Jones, M Hill, P Wilkes, C Axelsson, A Haywood, A Mellor. Searchingfor the Optimal Sampling Design for Measuring LAI in an Upland Rainforest. Proceedings of the Geospatial ScienceResearch Symposium GSR2, December 2012, Melbourne. ISBN: 978-0-9872527-1-5.

Acronyms

AAS/AATSE	Australian Academy of Science and the Australian Academy of Technological Sciences and Engineering
ABARES	Australian Bureau of Agricultural and Resource Economics and Sciences
Cal/Val	Calibration and Validation
CC	Canopy Cover
CEOS	Committee on Earth Observing Satellites
CEOS WGCV	Committee on Earth Observing Satellites Working Group on Calibration and validation
DHP	Digital Hemispherical Photography
ECV	Essestial Climate Variable
EO	Earth Observation
ESU	Elementary Sampling Unit
fAPAR	Fraction of Absorbed Photosynthetic Active Radiation
FPC	Foliage Projective Cover
fCover	Fractional Cover
FVC	Fractional Vegetation Cover
LAI	Leaf Area Index
LPV	Land Product Validation
LTER	Long Term Ecological Research Site
MODIS	Moderate Resolution Imaging Spectroradiometer

MODLAND	MODIS Land Discipline Team
NCAS	National Carbon Accounting System
NDVI	Normalized Difference Vegetation Index
NPP	Net Primary Production
SBA	Stand Basal Area
SSU	Secondary Sampling Unit
SLATS	State Land and Tree Survey
VALERI	Validation of Land European Remote sensing Instruments
WGCV	Working Group on Calibration and validation