On the Validation of Satellite-derived Products for Land Applications

Review Paper/Article de synthèse

by J. Cihlar • J. Chen • Z. Li

RÉSUMÉ

La validation rigoureuse de la précision et de la cohérence des produits dérivés des données satellitaires représentant des variables biophysiques de surface est essentielle au progrès de l'observation de la Terre en tant qu'outil d'analyse. La validation permet d'assurer que les produits dérivés rencontrent les caractéristiques techniques décrites et garantit que les résultats générés à l'aide de leur utilisation peuvent être considérés comme fiables. À titre de composante d'une approche globale de contrôle de la qualité, pratique d'usage courant dans la plupart des procédés de production, la validation est particulièrement importante pour les produits dérivés des données de télédétection parce que l'information est déduite et que le rapport du signal sur le bruit enregistré lors des mesures est fréquemment bas. Dans cet article, la validation des produits de télédétection est analysée en deux phases : la validation initiale qui permet d'établir la fiabilité d'un produit et la validation continue qui permet d'assurer la qualité et la fiabilité à plus long terme du produit. La discussion est axée sur les données satellitaires optiques et leurs applications terrestres, principalement dans l'environnement boréal. On y présente des exemples de produits d'applications au sol dérivés de données satellitaires, incluant un aperçu des difficultés réelles rencontrées lors de la validation initiale. Plusieurs stratégies sont proposées pour rendre la procédure de validation plus performante et efficace.

SUMMARY

Rigorous validation of the accuracy and consistency of satellite-derived products representing surface biophysical variables is at the heart of the success of Earth observation as a practical tool. Such validation ensures that the derived products meet the claimed specifications and thus results produced with their use can be viewed with confidence. As part of an overall quality control, which is involved in most production processes, validation is especially important for products derived from remote sensing data because the information is inferred and because the signal to noise ratio in the measurements is frequently low. In this paper, the validation of remote sensing products is considered in two phases: initial validation which establishes the soundness of a product; and continuing validation which ensures the ongoing quality and reliability of such a product. The discussion is focused on optical satellite data and their terrestrial applications, primarily in the boreal environment. Examples are given for land products derived from satellite data, including the practical difficulties in the initial validation. Several strategies are suggested to make the validation process more efficient and effective.

INTRODUCTION

The generation of information products from raw satellite measurements is difficult due to the complexity of the three-dimensional, often time-varying target medium; the variety of information sought from a limited number of measurement types; and the often tenuous relationship between the signal and the phenomenon of interest on one hand and the strength of the interfering effects on the other. Nevertheless, the development and implementation of the procedures for routine extraction of the information of interest as 'standard' products is critical to the expansion of the client community for the Earth observation technology, beyond the relatively limited number of remote sensing specialists who now constitute the bulk of the users. Importantly, the products must be of known quality, with accuracy information attached, so that the potential users can make informed decisions about the relevance of the product for their purpose.

The transformation of raw satellite measurements into an accurate biophysical or geophysical product involves a number of steps (**Figure 1**). The first stage is the conversion of the raw data into radiance or reflectance so that it represents the signal just after leaving the target of interest. This involves sensor calibration and atmospheric corrections. The data are then transformed into a product which represents a biophysical parameter describing the state or process in the Earth-atmosphere system. The transformation relies on an algorithm

• J. Cihlar, J. Chen and Z. Li are with the Canada Centre for Remote Sensing, Ottawa, Ontario K1A 0Y7

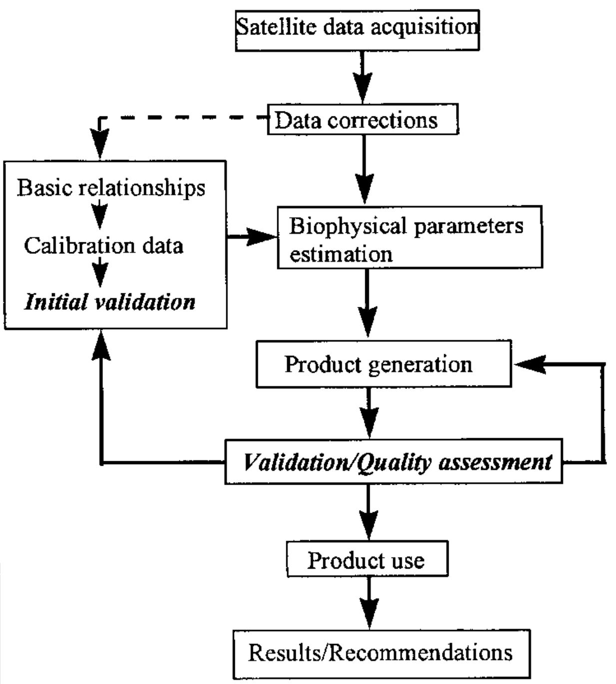

Figure 1.
Product validation in relation to the data handling stream from raw data to the use of the product.

or a model, usually derived from field or laboratory measurements, computer simulation, or a combination of these methods. To be considered for routine use, the algorithm itself must first be extensively tested and its accuracy and robustness proven as part of the initial algorithm definition and development process; we call this step 'initial validation' (**Figure 1**). Initial validation involves the study of the basic relationships between the reflected or emitted radiation and the variable of interest. It may require laboratory studies of this relationship; field measurements under a range of environmental conditions; coincident field, airborne and/or satellite measurements; modelling of the relationship between the variable of interest and the reflected or emitted radiation with various assumptions and degrees of complexity; or, more typically, a combination of these approaches. During initial validation, sets of calibration data are produced for use in establishing the feasibility and accuracy of the retrieval of the variable of interest from remotely sensed data. When completed, the initial validation yields an algorithm for retrieving the variable of interest from remotely sensed data, together with the detailed specifications of the retrieval procedures and parameters to be used, the accuracy, and error analysis.

Once a product is routinely generated from the satellite data using the above algorithm, further validation is required as part of quality assessment (**Figure 1**). The second validation step assesses whether the product continues to reflect reality (as described by independent observations). The two-step validation is necessary to avoid generating unsatisfactory products even with a good algorithm. In this context it is useful to distinguish various levels of products. Using the nomenclature defined by the Committee on Earth Observation Satellites (CEOS), three product levels are relevant to the discussion: level 2 (retrieved environmental variables at the same resolution and location as level 1 data, i.e. calibrated data in satellite projection); level 3 (retrieved environmental variables that have been spatially and/or temporally resampled); and level 4 (model output or results from analyses of lower level data). Since all three levels have a strong remote sensing heritage, albeit to various degrees, the Earth observation community needs to be concerned with the validation at all three levels. In this paper, we consider some issues related to both validation stages and specifically those associated with the quality of the final information products about the land surface.

The challenge of satellite product validation can thus be expressed by the following question: how closely does the product correspond to the same variable in reality, in both magnitude and distribution? Using the above terminology it includes both initial and continuing validation; the various steps included in the two phases, from basic research on the radiation – target interactions to operational satellite data acquisition, processing and product generation; and the assumptions, uncertainties and other limitations or errors in this process. The type and relative importance of these factors may vary with product and application. A comprehensive, detailed analysis of the various steps is beyond the scope of this paper. Thus, we address some generic and some specific issues based on our research experience, mainly in relation to biophysical land variables and the boreal environment.

Product Validation

Initial validation

By initial validation we mean the process of establishing the quality of an algorithm by assessing the product generated by the algorithm. Several approaches are possible.

(i) A common method is the comparison of the product with an independently obtained data set representing the same information (time/space/variable domain). Such data sets are obtained through a variety of sources, most frequently in field data collection campaigns. If products are validated with independent data sets, it is implicitly assumed that the latter are a 100% accurate representation of reality. In practice, that is rarely the case as data collected *in situ* usually have errors and may be biased through sampling, measurement methods, etc. Some errors are random, caused by noise from many unwanted instrumental and environmental variations, and can be largely reduced by multiple or repeat measurements. Others are bias errors that cannot be removed after data averaging operations. Bias errors can usually be found using independent or improved techniques but they are difficult to assess unless such techniques are available. Note that geophysical measurements may be correlated in time and space if they are collected continuously and consecutively, thus violating the independence assumption of classical statistical analysis (e.g., Li and Leighton, 1992).

(ii) Another validation approach is to compare independently derived (data source and algorithm) products from satellite data sets, i.e. measuring the same variables in different ways. The comparison is strengthened if one set has already been validated against ground data. For example, Li (1995) compared two satellite-based surface radiation budget products. Weaknesses and strengths of both products were identified by analyzing the differences caused by the different input data and algorithms. The analysis of discrepancies between the two data sets may thus yield new understanding of the algorithms or of data from which the products were derived, thus giving an opportunity for further improvements. When interpreting the results of inter-comparisons, it should be kept in mind that the agreement between two independent products does not necessarily confirm the accuracy of either. It is not uncommon that both products are subject to the same deficiencies or bias errors or that errors caused by some factors may be canceled out by others, leading to a false agreement.

(iii) Alternatively, the derived products may be compared to outputs of physically based models that describe the underlying physical and biophysical processes governing remote sensing signals from the target of interest. Frequently, the algorithm used to generate the product itself results from a simplification of a more complex model, after exercising the latter under a variety of conditions. Models of this type can be extrapolated to areas where field data are not available and the model assumptions remain valid. Models thus greatly enhance the usefulness of limited data sets and their potential should not be underestimated. However, they also have limitations.

(1) They are always to some degree simplified mathematical descriptions of reality. Details ignored by models that are insignificant in some cases can become important in other cases.
(2) In formulating mathematical descriptions and model simplifications, assumptions have to be made. The assumptions are usually subject to the level of understanding of the modeled process and to the tools and facilities available to that modeler. In making predictions or extrapolations using models, the results are reliable only if the assumptions remain valid. In reality, assumptions are usually violated to some extent, incurring errors.
(3) Physically-based models often require input variables that cannot be obtained remotely. These variables are sometimes acquired by "tuning" models to produce the desired output. Retuning a model is often needed when applied to a new area. This problem of intrinsic remote sensing dimensionality being smaller than the required modeling dimensionality imposes a serious limitation to the usefulness of many models (Hall *et al.*, 1995b).
(4) Computationally, models are usually further simplified for applications over large areas, introducing additional uncertainties in derived products.

For the above reasons, an optimum approach to product validation is to use both field data and models in a parallel and interactive manner (Hall *et al.*, 1992). Models need to be validated with some ground-truth data, but in turn they can provide guidance for improving field experimental strategy and extending the validation. For example, further collection of field data can be directed toward testing some crucial assumptions made in models. Models may also reveal inadequacies of input data, e.g. sampling or measurement biases.

Continuing validation

Once algorithms for the conversion of data into information products are developed and proven to perform well, several questions remain:

- How well does the algorithm perform in the same area but another year?
- How well does the algorithm perform in another geographic area, in the same or somewhat different ecological setting?
- How well does the algorithm perform with different satellite data sets?

If the characteristics of the input data do not change, the product should continue to retain its accuracy over extended periods. Thus, the principal reason for concern is the fact that in practice, the input data are not likely to remain static even for a given sensor, as the sensor response tends to vary with time. The calibration of input data and their stability with time therefore become critical for the performance of the algorithm used and the consistency of the final product. Continuity of data sets becomes a critical issue when data from various sensors are to be used to produce a time series of a product. Such series may consist of the same sensor type flown on successive platforms or of different sensors making similar measurements. For example, data from AVHRR onboard NOAA-7,-9,-11,-14; EOS/MODIS; and SPOT/VEGETATION will form a time series for the assessment of global vegetation dynamics over a 25-year period. How can the continuity and consistency of products such as FPAR (fraction of photosynthetically active radiation) and LAI (leaf area index) be assured? Unless suitable historical data sets exist, an independent assessment of retrospective products becomes impossible. The most powerful tool in this case is probably a comparison of the product for overlapping periods when data from more than one sensor are obtained. When the compared sensors produce the same results one is reassured that the product is likely satisfactory. However, if discrepancies are found, it may not be clear why they occur and which product is closer to the actual values. In any case, the continuity of the calibrated measurements is critical and an assumption must be made that the fundamental relationship between the remote sensing measurements and the biophysical variables of interest has not changed.

To ensure the stability of information products derived from satellite data, a strong need exists for a continuing or a periodic independent data set with which the remote sensing products may be compared. This is the case even in measurements where the signal is quite strong, but it becomes critical when the signal is mixed with various sources of noise, as for example the information about green vegetation phytomass in terrestrial ecosystems. For land applications, networks that provide ground data for periodic validation do not presently exist. Although environmental variables (species composition, canopy structure, leaf area index) are measured at some sites in various countries, the measured variables are rarely obtained in a way that yields a spatially and thematically representative and independent data set. This situation has led to discussions about a system for long-term observations of land variables that could be used in conjunction with satellite observation to detect global change trends (Global Climate Observing System, 1995). It should be noted that an independent data set obtained to assess the accuracy of satellite-derived products must meet certain criteria of representativeness, location specificity and accuracy that are, at least in some respects, likely to exceed the specifications for the same variables collected traditionally or for other reasons.

Examples

The validation issues for land surface variables are illustrated below for three variables: boreal forest leaf area index, surface solar radiation budget, and absorbed photosynthetically active radiation.

Boreal forest leaf area index

Although boreal forests often consist of single species in fairly even age stands that have simpler foliage structure than tropical and temperate forests, they also present a challenge to assessing the leaf area index (LAI) over large areas. The major problems in the two-stage validation process for LAI algorithms include the following (the first six concern initial validation, the last one relates to continuing validation):

(1) Ground data collection. Because of the labour involved and the impact on the forest, only limited destructive sampling of LAI is possible. Therefore, indirect methods are the principal means for obtaining actual leaf area data. Deriving LAI based on allometric methods (relating easily measured stand variables to leaf area) can cause large errors because stand conditions change considerably across the landscape. Allometric relationships are often stand specific and depend on species, age, stand density and other conditions (Smith, 1993). They require destructive sampling which is laborious and usually involves substantial errors (Chen, 1996a).

Since boreal forest canopies, especially conifers, have a highly organized architecture, LAI measurements by many existing optical instruments are biased if the spatial distribution of foliage is assumed to be random. Nevertheless, optical methods can be more accurate than allometric methods if the effect of foliage architecture on LAI measurements can be estimated. Chen and Cihlar (1995a) developed a sunfleck-LAI instrument called TRAC (Tracing Radiation and Architecture of Canopies) to measure the effect of foliage clumping at scales larger than shoots based on a new canopy gap size analysis theory (Chen and Cihlar, 1995a). Although optical measurements of LAI have been substantially improved with the use of TRAC, the problem with foliage clumping within shoots still remains. Techniques for determining this small-scale clumping effect are available (Fassnacht *et al.*, 1994; Chen, 1996a) but a large number of shoots need to be processed to characterize a stand. Chen (1996a) determined this within-shoot clumping effect for six boreal conifer stands at three stages during the growing season and found that the effect varies about 15% between species and ages. It appears that combined use of the LI-COR LAI-2000 (for foliage angle distribution) and the TRAC (for foliage spatial distribution) can allow rapid estimation of LAI for large areas, with an accuracy comparable to ground sampling (20 - 30%).

(2) Choice of vegetation indices. For LAI mapping, vegetation indices remain simple and effective tools although alternatives have been developed using simple inversion models (Kuusk, 1991; Hall *et al.*, 1995a). Chen (1996b) found that near infrared (NIR) reflectance is insensitive to the changes in LAI of boreal forest stands because the increased reflection by green leaves with increasing LAI is counter-balanced by the increased tree crown shadows, which have low reflectance. The major signal of LAI lies in the red band where an increase in LAI causes monotonic decrease in reflectance; this is because the shadow effect augments the decrease in leaf reflection with increasing LAI. The small sensitivity of NIR reflectance to LAI disqualifies those vegetation indices designed to maximize the information content in NIR measurements. For LAI mapping, the NIR channel is most effectively used in indices based on the ratio of NIR to red reflectance, such as the Simple Ratio (SR) and the Normalized Difference Vegetation Index (NDVI) because the ratio reduces many unwanted sources of environmental noise in both channels (**Figure 2**). There are many sources of environmental noise, including subpixel clouds and shadows, dissimilar surface features, subpixel water bodies, edge effects, etc. All these sources of noise generally cause simultaneous increase or decrease in the reflectance for both red and NIR channels in about the same proportion, and they are largely reduced when the ratio between these channels is taken. In the boreal forest environment where the background is usually composed of green moss and understory which differ relatively little from the overstory greenness, the noise reduction mechanism of ratio-based indices (NDVI, SR) is extremely important in the success of LAI mapping using remote sensing measurements. Chen (1996b) showed that the SAVI (Huete, 1988) and other indices formulated for suppressing the effect of background on vegetation information retrieval are not suitable for the boreal environment because they are not in accord with the ratio principle (**Figure 2**). Chen (1996b) also showed that better linear relationships exist between SR and LAI than between NDVI and LAI. The usefulness of SR for mapping LAI in temperate forests is shown in the recent study by Fassnacht *et al.* (1997). The mid-infrared band can also be used to modify SR for improving LAI derivation (Nemani *et al.*, 1993).

(3) Species dependence. Chen and Cihlar (1996) correlated ground- based LAI measurements with Landsat 5 TM NDVI and found only small differences between black spruce and jack pine species. However, the small sample size and noise in the data made it difficult to assess if the difference is significant.

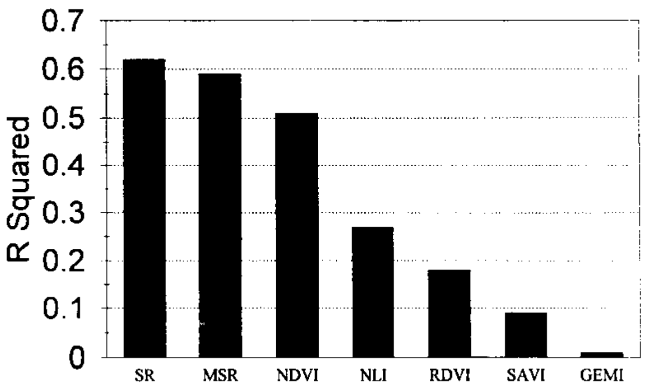

Figure 2.
The coefficient of determination (r^2) between two-band vegetation indices derived from Landsat TM and LAI values measured in 20 conifer (jack pine and black spruce) stands in Saskatchewan and Manitoba in late spring 1994. The indices include the Simple Ratio (SR), Modified Simple Ratio (MSR), Normalized Difference Vegetation Index (NDVI), Non-Linear Index (NLI), Renormalized Difference Vegetation Index (RDVI), Soil Adjusted Vegetation Index (SAVI), and Global Environmental Monitoring Index (GEMI). Refer to Chen (1996b) for more detail.

The background (moss cover and understorey) spectral measurements from these two types of stands were significantly different (Miller *et al.*, 1997). The optical properties of jack pine and black spruce needles were also found to be different (Middleton *et al.*, 1997). These recent results indicate that it might be necessary to distinguish between these two species in LAI algorithm development. The NDVI and other vegetation indices are considerably greater over deciduous forests (aspen) than over coniferous forests (Loechel *et al.*, 1997).This suggests that for LAI estimation, a corresponding forest cover map with classes of conifer, deciduous and mixed forests (as a minimum) is necessary.

(4) Understory effect. Boreal forests are open with considerable understory growth. The amount and species of understory vary depending on the density and species of the overstory. Many understory species in the boreal forest (including alder, labrador tea, cranberry, blueberry, sphagnum moss, feather moss) have NDVI values that are only 0.15 and 0.25 smaller than those of overstory vegetation (White *et al.*, 1995). This makes it difficult to determine the overstory LAI. This problem is reduced when NDVI is measured in early spring before the understory growth occurs (Chen and Cihlar, 1996). However, to achieve an accuracy of LAI within 25%, NDVI should be known to within 0.05. This implies the need for very accurate solar zenith angle, view geometry and atmospheric corrections and very accurate sensor-to-sensor calibrations when algorithms derived for one satellite sensor are used for other sensors.

(5) Scale effect. Boreal landscape is heterogeneous with a large fraction of open water. The large contrast between water and forest poses a problem for moderate and coarse resolution images. Large pixels frequently contain (often small) open water bodies that introduce a large bias in LAI estimation based on algorithms derived and validated using high resolution data. After an intensive study on the effect of small water bodies we concluded that a correction factor can be applied for LAI estimation based on percent water cover in the pixel. This suggests that for accurate determination of biophysical variables over large areas, an open water mask at a higher spatial resolution is required (Chen, 1998).

(6) Temporal stability. Vegetation indices of both deciduous and coniferous boreal forests exhibit strong seasonal dynamics. Most boreal conifer trees carry the same needles for up to 4 - 10 years. New needles during the growing season contribute to about 25 - 30% of the total foliage area (Chen, 1996a). However, NDVI from conifer stands has much larger seasonal variability (Cihlar *et al.*, 1997a). This may be due to the understory condition and the effects of seasonal changes on the chlorophyll content. The temporal variability should be considered in developing algorithms for both coniferous and deciduous forests. The dependence of NDVI on solar zenith angle should also be considered when images from different seasons and latitudes are compared (e.g. Hall *et al.*, 1995), or when measurements made at the same location but at different times are compared (Leblanc *et al.*, 1997).

(7) Change of satellite data sources. Because of the high sensitivity of estimated LAI to vegetation indices, algorithms are generally sensor-specific, i.e. a small difference in NDVI due to sensor calibration or spectral bandwidth choices can cause large errors in the derived products (Teillet *et al.*, 1997). Careful sensor cross-calibration is needed when algorithms developed for one sensor are used for data from other sensors.

Surface solar radiation budget

While numerous attempts have been made to retrieve and validate surface solar radiation budget (SRB) (Gautier *et al.*, 1980; Pinker and Laszlo, 1992; and Darnell *et al.*, 1992; among others), the discussion here is concentrated on the work of Li and his colleagues who examined several aspects of product validation.

Li *et al.* (1993a) developed an algorithm for retrieving the SRB from top-of-the-atmosphere (TOA) reflected solar radiation measured with satellite sensors. Since the algorithm is based exclusively on radiative transfer modeling, comprehensive initial validation was carried out using both ground-based observations and an independent data set. A direct validation was done by comparing co-located (in time and area) satellite-based estimates with surface measurements (Li *et al.*, 1993b). Although surface data were taken from a tower that increased the field-of-view (FOV) of the surface radiometer to better match satellite pixels, differences in space and time between the two types of data are still considerable, which obscures the actual discrepancies. Since such space-time differences are essentially random, the scatter of the comparison is generally much larger than the inherent random error of the satellite-based product. On the other hand, the mean difference between the satellite and surface values can represent the real bias error, provided that sufficient samples are available. This turns out to be close to zero under all-sky conditions (Li *et al.*, 1993b).

An indirect estimation of the random error embedded in the satellite-based SRB products took advantage of the variable density of global surface radiation network (Li *et al.*, 1995).

Since the standard deviation of the differences between observed and estimated grid-mean SRB decreased rapidly with the density of ground stations (**Figure 3**), the scatter of the points stems mainly from the time-space mismatch of the two data sets. By fitting the trend with an empirical function, one can isolate the real random error as the asymptotic value corresponding to infinite density of surface stations, which was found to be around 5 Wm^{-2} for the monthly mean SRB product (Li *et al.*, 1995).

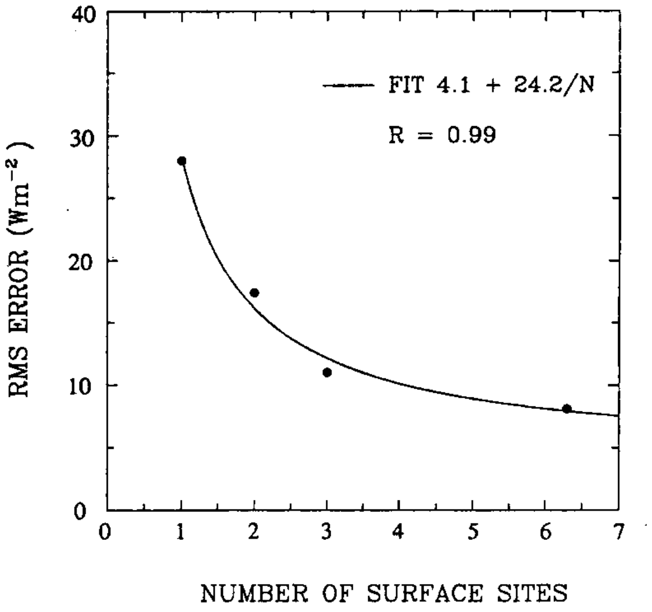

Figure 3.
The variation of the standard deviation of differences between observed and estimated surface solar radiation budget with the average number of surface observation stations within a grid cell of 280 km by 280 km. Surface measurements were taken from the global radiation networks, while satellite-based estimates were derived from the Earth Radiation Budget Experiment. Both data are for the period of 1985 to 1989.

In addition to the global comparisons described above, an analysis of the spatial and/or temporal pattern of the differences can also be helpful in revealing some characteristics of a remote sensing product. If the pattern resembles the distribution of a variable that is related to the derived parameter, the variable may be treated inadequately. For example, when the satellite SRB data were compared with global surface observations, large differences occurred over the tropical regions with active biomass burning (Whitlock *et al.*, 1996; Li, 1997). This revealed deficiencies in the algorithm's treatment of smoke-related aerosols.

As another example, consider surface albedo which is an important component of the surface solar radiation budget. Li and Garand (1994) developed an algorithm for retrieving clear-sky surface albedo from TOA satellite observed albedo. Global monthly mean surface albedos at a resolution of 2.5° by 2.5° were obtained by applying the algorithm to the Earth Radiation Budget Experiment data. Due to the lack of grid-mean ground-truth measurements, an assessment of this product was made by comparing with other satellite-based products (e.g., Staylor and Wilber, 1990). Unfortunately, large discrepancies appear to occur among the existing products, although Li and Garand's (1994) estimates agree with most of the data available. The algorithm was validated using tower measurements made in Boulder, Colorado, and Saskatoon, Saskatchewan. No bias and small random errors were found (Li and Garand, 1994).

Although the above tests conducted during initial validation confirm the adequacy of a particular data set, the longer-term problem regarding SRB measurements remains. The main satellite input for retrieving SRB and surface albedo is the TOA albedo or reflected flux that is derived from TOA radiance measurements after radiometric calibration and bidirectional correction. Both corrections are sensor-dependent and therefore the quality of the retrieved surface products may vary from one sensor to another, unless the corrections are free from uncertainties or the uncertainties remain constant. Since Li and Garand's (1994) algorithm is fully based on radiative transfer theory, its performance has minimal regional dependence, but the quality of input data may vary considerably from one region/time to another and so could the output data. Therefore, long-term observations from a widely distributed surface radiation network such as the Baseline Surface Radiation Network are essential for the continuing validation of global long-term SRB products.

Canopy absorbed photosynthetically active radiation (APAR)

Photosynthetically active radiation (PAR) refers to the solar radiation in the spectral range 400-700 nm. A new technique for remote sensing of the PAR absorbed by the green foliage of the canopy, APAR, was proposed (Li and Moreau, 1996) that first estimates the amount of radiation absorbed by all materials below the top of the canopy (APART). APAR is then computed from APART, the fraction of the PAR absorbed by green foliage (FPAR), and PAR surface albedo. The estimation of APART employs a physically-based algorithm that requires the inputs of upwelling PAR at the TOA, ozone amount, and aerosol optical thickness and its absorbing property (Li and Moreau, 1996). Upwelling PAR at the TOA is the most important variable whose changes largely drive the variation of APART. Surrogate measurements of upwelling PAR can be obtained from the visible channels of many existing and future satellite sensors. The accuracy of information on aerosol is of secondary importance (Li and Moreau, 1996).

Similar to the validation of SRB, matching surface and space-borne observations is the most critical aspect in the initial validation of the APART accuracy. Unfortunately, very few upwelling and downwelling PAR measurements made at the top of the canopy, and even fewer co-located with satellite observations. To date, PAR measurements have been made almost exclusively in field experiments of short duration, such as FIFE and BOREAS. Preliminary validations were done using these field measurements (Li and Moreau, 1996; Li *et al.*, 1997).

The fraction of APART absorbed by the canopy can be determined following the method described by Moreau and Li

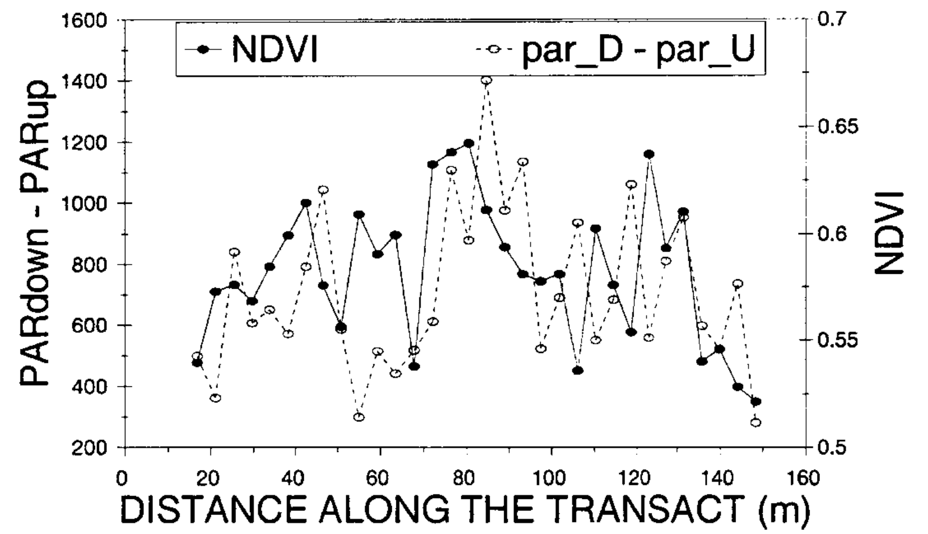

Figure 4.
A comparison of the variations in understory net PAR (down minus up) measured below tree crown with the TRAC and the Normalized Difference Vegetation Index (NDVI) measured above the canopy with the CASI airborne sensor. The observations were taken at a young jack pine site in the BOREAS southern study area during 1994. The two types of data sets were co-registered.

(1996) or from FPAR. Initial validation of these methods requires measurements of downwelling and upwelling PAR fluxes at the top and bottom of canopies, together with information pertaining to the greenness of the canopy under study (e.g., using NDVI). Surface NDVI can be derived from satellite visible and near-infrared measurements after atmospheric corrections, or directly from near-surface airborne data. Validation of the proportion of PAR absorbed by canopies is thus much more difficult than of APART. Not only are there few coincident *in situ* and remote measurements available, these measurements also suffer considerable differences in spatial representation. The problem is further hampered by the variability of vegetation coverage in the scale comparable to the size of a tree crown. To resolve such local variability, data collected with the TRAC instrument (Chen, 1996a) over an area or along a transect are needed. TRAC uses walking and high-frequency sampling techniques with upward- and downward-facing PAR sensors to obtain measurements at a spatial interval of 10 mm along 150-350 m transects beneath the canopy, about 0.5 m above the soil surface.

As an example of the complexity of the initial validation of FPAR, **Figure 4** shows the correspondence between two data sets, airborne and below-canopy. CASI airborne data obtained during BOREAS were processed to compute NDVI at a resolution of 3m by 3m (Miller *et al.*, 1995) and compared with net PAR flux measured below canopy with the TRAC. Since the above canopy PAR is much less variable over short distances, the fluctuations in **Figure 4** reflect mainly the below-canopy PAR difference (down minus up) in relation to NDVI. In many instances the two variables change out of phase, i.e. larger PAR differences correspond to smaller NDVI. This is understandable as smaller NDVI implies fewer leaves, allowing more PAR photons to pass through the canopy. However, this relation is by no means universal as indicated by the nearly in-phase variation in the middle part of the transect. In fact, **Figure 4** is among the very few "good" examples from many similar analyses. There are numerous factors that can distort the correspondence between the two variables, including a misregistration of the two data sets and the difference in the scenes sensitive to below and above-canopy observation. PAR transmittance measured below the canopy is more sensitive to the tree crown along the pathway of direct solar beam, whereas the NDVI measurements are sensitive to all scene elements visible from above within the field of view. For example, understory vegetation has little impact on downwelling PAR, but it may dictate NDVI entirely if there are no trees nearby. Under such circumstances, one cannot expect any correlation between the two quantities. Because of these and other factors, validation of FPAR and other vegetation attributes inferred from remote sensing is a complex undertaking that involves numerous inherent uncertainties.

VALIDATION STRATEGIES

The basic strategy for the initial validation has been the use of comprehensive field campaigns. These have been carried out for diverse land, ocean, and atmospheric applications. For land, they have been conducted in various biomes: grassland (Sellers *et al.*, 1992), grassland/savannah (Prince *et al.*, 1995), boreal forest (Sellers *et al.*, 1995), and others. Algorithm development and initial validation, supported by field measurements, are major objectives in these experiments. These experiments often generate a number of remote sensing data sets, algorithms and products with the initial validation results (e.g., Cihlar *et al.*, 1997b).

Ongoing product validation implies an approach in which the characteristics of the generated product are systematically assessed against 'expectations' and any deviations are flagged for more detailed analysis. Such validation must be built into the product generation cycle. The 'expectations' refer to the expected values of the geophysical product, which may be known only for a sample of the domain, at best. This sample can be produced from a set of reference sites where the parameter is routinely measured. Various qualitative assessments are also possible, e.g. the values should not exceed certain thresholds under given conditions. However, the overall accuracy of the product is very difficult to ascertain. The process can be expensive and time-consuming, but also as intellectually stimulating as the development of new models or discovery of new relationships in experimental data. It is a critical and essential ingredient in ensuring the success of Earth observation science and technology from space. Below are some suggested approaches that should reduce the effort needed to produce consistent time series of derived products. Some of these have been illustrated in the preceding sections. They are divided into two groups, those relevant to validation strategies and those aimed at obtaining appropriate remotely sensed data sets.

For research:

• Use algorithms that scale easily. For example, algorithms based on linear relationships between satellite measurements and the geophysical variable can be transferred much more readily from high resolution (e.g. TM) to lower resolution (AVHRR, MODIS) data since only the spectral equivalence needs to be formulated. Such products should be easier to validate if produced by sensors with different spatial resolutions.

• Use other sensors as a bootstrap to span discontinuities. For example, AVHRR pixels may be characterized using TM data, in turn described with ground data obtained directly.

• Use products derived independently from various input data/sensors to assess consistency of results.

• Use model sensitivity tests to determine whether particular environmental variables are important and should be included in ground data collection and algorithm validation efforts.

• Use various models (developed independently with different assumptions) as an interim step in validation.

• Results of the applications of the data in models, i.e. higher level products, may also provide useful input into validation. This is because the deficiencies may not become evident in the product itself but will appear in the final result because of the amplification of the errors by the additional processing or because of other inputs.

Regarding remote sensing data:

• Overlap operations of different sensors so that products can be compared for the common period. It is important to include a representative range of conditions in the comparison.

• When field data collection activities are undertaken, remote sensing data from as many potentially applicable sensors as possible should also be obtained, together with measurements of all algorithm input and output variables.

The above approaches are not mutually exclusive, and in fact many can be used as part of the validation scheme for a particular product. As satellite remote sensing becomes increasingly more relied on to deliver quantitative information such validation strategies must be incorporated as an integral component of the product generation cycle.

ACKNOWLEDGMENTS

We gratefully acknowledge the provision of the CASI data used to produce results in **Figure 4** by the York University team led by Dr. John Miller. Dr. Philippe Teillet made helpful comments on a draft version of this paper.

REFERENCES

Chen, J.M. 1996a. "Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands". *Agricultural and Forest Meteorology* 80: 135-163.

Chen, J.M. 1996b. "Evaluation of vegetation indices and a Modified Simple Ratio for boreal applications". *Canadian Journal for Remote Sensing* 22: 229-242.

Chen, J. M. 1996c. "Canopy architecture and remote sensing of the fraction photosynthetically active radiation absorbed by boreal conifer forests". *IEEE TransactionsTrans on Geoscience and Remote Sensing*, 34: 1353-1368.

Chen, J. M., 1998. "Quantifying spatial scaling effects on remote sensing of surface parameters". *Remote Sensing of Environment* (in press).

Chen, J.M., and J. Cihlar. 1995a. "Plant canopy gap size analysis theory for improving optical measurements of leaf area index of plant canopies". *Applied Optics* 34: 6211-6222.

Chen, J.M., and J. Cihlar. 1996. "Retrieving leaf area index of boreal forests using Landsat TM images". *Remote Sensing of Environment* 55: 153-162.

Chen, J. M., P. M. Rich, T. S. Gower, J. M. Norman, S. Plummer, 1997. "Leaf area index of boreal forests: theory, techniques and measurements". *Journal of Geophysical Research.* (in press).

Cihlar, J., J. Chen, Z. Li, F. Huang, and H. Pokrant. 1997a. "How well can interannual land surface signal be discerned in noisy AVHRR data?" *Journal of Geophysical Research* (under review).

Cihlar, J., J. Chen, and Z. Li. 1997b. "Seasonal AVHRR multichannel data sets and products for studies of surface-atmosphere interactions". *Journal of Geophysical Research.* (in print)

Darnell, W.L., W.F. Staylor, S.K. Gupta, N.A. Ritchey, and A.C. Wilber. 1992. "Seasonal variation of surface radiation budget derived from International Satellite Cloud Climatology Project C1 data". *Journal of Geophysical Research* 97: 15741-15760.

Fassnacht, K., S.T. Gower, J.M. Norman, and R.E. McMurtrie. 1994. "A comparison of optical and direct methods for estimating foliage surface area index in forests", *Agricultural and Forest Meteorology*. 71: 183-207.

Fassnacht, K. S., S. T. Gower, M. D. MacKenzie, E. V. Nordeim, and T. M. Lillesand, 1997. "Estimating the leaf area index of north central Wisconsin forests using the Landsat Thematic Mapper", *Remote Sensing of Environment* 61:229-245.

Frouin, R., and R.T. Pinker. 1995. "Estimating photosynthetically active radiation (PAR) at the Earth's surface from satellite observations". *Remote Sensing of Environment* 51:98-107.

Frouin, R., and C. Gautier. 1990. "Variability of photosynthetically available and total solar irradiance at the surface during FIFE: A satellite description". *Proceedings of the Symposium on the First ISLSCP Field Experiment*, Feb. 7-9, 1990, Anaheim, California, American Meteorological Society, Boston, Massachusetts: 98-104.

Gautier, C., G. Diak, and S. Masse. 1980. "A simple physical model to estimate incident solar radiation at the surface from GOES satellite data". *Journal of Applied Meteorology* 19: 1005-1012.

Global Climate Observing System. 1995. *Plan for the Global Climate Observing System (GCOS)*. Report GCOS-14, WMO/TD - No. 681, World Meteorological Organization. 49p.

Gu, J., and E. Smith. 1997. "High resolution estimates of total solar and PAR surface fluxes over large scale BOREAS study area from GOES measurements". *Journal of Geophysical Research* (in press).

Hall, F.G., K.F. Huemmrich, S.J. Goetz, P.J. Sellers, and J.E. Nickeson. 1992. "Satellite remote sensing of surface energy balance: success, failures, and unresolved issues in FIFE". *Journal of Geophysical Research* 97 (D17): 19061-19089.

Hall, F.G., Y.E. Shimabukuro, and K.F. Huemmrich. 1995a. "Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models". *Ecological Applications* 5: 993-1013.

Hall, F. G., J.R. Townshend, and E.T. Engman. 1995b. "Status of remote sensing algorithms for estimation of land surface state parameters". *Remote Sensing of Environment* 51:138 - 156.

Huete, A.R. 1988. "A soil adjusted vegetation index (SAVI)". *Remote Sensing of Environment* 25:295-309.

Kuusk, A. 1991. "Determination of vegetation canopy parameters from optical measurements". *Remote Sensing of Environment* 37: 207-218.

Leblanc, S., J. M. Chen, and J. Cihlar. 1997. "Directionality of the NDVI in boreal forests". *Canadian Journal for Remote Sensing* (in press, this issue).

Li, Z., and H.G. Leighton. 1992. "Narrowband to broadband conversion with spatially autocorrelated reflectance measurements". *Journal of Applied Meteorology* 31: 421-432.

Li, Z., H.G. Leighton, K. Masuda, and T. Takashima. 1993a. "Estimation of SW flux absorbed at the surface from TOA reflected flux". *Journal of Climate* 6: 317-330.

Li, Z., H.G. Leighton, and R.D. Cess. 1993b. "Surface net solar radiation estimated from satellite measurements: Comparisons with tower observations". *Journal of Climate 6*: 1764-1772.

Li, Z., and L. Garand. 1994. "Estimation of surface albedo from space: A parameterization for global application". *Journal of Geophysical Research* 99: 8335-8350.

Li, Z., C.H. Whitlock, and T.P. Charlock. 1995. "Assessment of the global monthly mean surface insolation estimated from satellite measurements using global energy balance archive data". *Journal of Climate* 8: 315-328.

Li, Z., 1995. "Intercomparison between two satellite-based products of net surface shortwave radiation". *Journal of Geophysical Research* 100: 3221-3232.

Li, Z.. 1998. "Influence of absorbing aerosols on the solar surface radiation budget". *Journal of Climate* 11 (in press).

Li, Z., and L. Moreau. 1996. "A new approach for remote sensing of canopy absorbed photosynthetically active radiation, I: Total Surface Absorption". *Remote Sensing of Environment* 55: 175-191.

Li, Z., L. Moreau, and J. Cihlar. 1997. "Estimation of the photosynthetically active radiation absorbed at the surface over the BOREAS region". *Journal of Geophysical Research.* (in press)

Loechel, S. E., C. L. Walthall, E. Brown de Colstoun, J. M. Chen, B. L. Markham, and J. Miller, 1997. "Model-based analysis of canopy and understory reflectances at BOREAS". *Journal of Geophysical Research* (in press).

Middleton, E. M., S. S. Chan, R. J. Rusin, and S. K. Mitchell, 1997. "Optical properties of black spruce and jack pine needles at BOREAS sites in Saskatchewan, Canada". *Canadian Journal of Remote Sensing*, 23:108-119.

Miller, J., J. Freemantle, P. Shepherd, L. Gray, N. O'Neill, A. Royer, and E. Senese. 1995. "Deployment of CASI to meet the needs of BOREAS science". *Proceedings of the 17th Canadian Symposium on Remote Sensing*, 13-17 June, 1995, Saskatoon, Saskatchewan: 169-175.

Miller, J. R., H. P. White, J. M. Chen, G. McDermid, D. Peddle, R. Fournier, P. Shepherd, I. Rubinstein, J. Freemantle, R. Stouffer, and E. LeDrew. 1997. "Seasonal change in the mean understory reflectance at BOREAS flux-tower sites and its application to canopy vegetation index determination". *Journal of Geophysical Research.* (in press).

Moreau, L., and Z. Li. 1996. "A new approach for remote sensing of canopy absorbed photosynthetically active radiation. II: Proportion of canopy absorption". *Remote Sensing of Environment* 55: 192-204.

Nemani, R., L. Pierce, S. Running, and L. Band. 1993. "Forest ecosystem processes at the watershed scale: sensitivity to remotely-sensed leaf area index estimates". *International Journal of Remote Sensing* 14:2519-2534.

Pinker, R.T., and I. Laszlo. 1992. "Modeling surface solar irradiance for satellite applications on global scale". *Journal of Applied Meteorology* 31:194-211.

Prince, SD., Y.H. Kerr, J.-P. Goutorbe, T. Lebel, A. Tinga, *et al.* 1995. "Geographical, biological and remote sensing aspects of the Hydrologic Atmospheric Pilot Experiment in the Sahel (HAPEX-Sahel)". *Remote Sensing of Environment* 51: 215-234.

Sellers, P.J., F.G. Hall, G. Asrar, D.E. Strebel, and R.E. Murphy. 1992. "An overview of the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE)". *Journal of Geophysical Research* 97 (D17): 18345-18371.

Sellers, P., F. Hall, H. Margolis, R. Kelly, D. Baldocchi, J. den Hartog, J. Cihlar, M. Ryan, B. Goodison, P. Crill, J. Ranson, D. Lettenmeier, and D. Wickland. 1995. "The Boreal Ecosystem-Atmosphere Study (BOREAS): an overview and early results from the 1994 field year". *Bulletin of the American Meteorological Society* 76: 1549-1577.

Smith, N. J. 1993. "Estimating leaf area index and light extinction coefficients in stands of Douglas fir". *Canadian Journal of Forest Research* 6: 49 57.

Staylor, W.F., and A.C. Wilber. 1990. "Global surface albedos estimated from ERBE data". *Proceedings of the 7th AMS Conference on Atmospheric Radiation*, San Francisco, California, American Meteorological Society, Boston, Massachusetts: 237-242.

Teillet, P.M., K. Staenz, and D.J. Williams. 1997. "Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indices of forested regions". *Remote Sensing of Environment* 61: 139-149.

White, H. P., J.R. Miller, J.M. Chen, D.R. Peddle, and G. McDermid. 1995. "Seasonal change in mean understory reflectance for boreal sites: preliminary results". *Proceedings of 17th Canadian Symposium on Remote Sensing.* Saskatoon, Saskatchewan: 182-187.

Whitlock, C.H., D.R. Cahoon, and T. Konzelmann. 1996. Biomass burning effects on satellite estimates of shortwave radiation in Africa. In: Levine, J.S. (Ed.), *Biomass Burning and Global Change*, The MIT Press: 485-491.