National Physical Laboratory

NPL REPORT ENV 5

Quality Assurance for FOREST data products: Recommendations for Future Improvements

Alexandra Barker
Joanne Nightingale
Andrew Banks
Tracy Scanlon

DECEMBER 2015

Quality Assurance for FOREST data products:
Recommendations for Future Improvements

Alexandra Barker, Joanne Nightingale, Andrew Banks and Tracy Scanlon
Earth Observation, Climate and Optical Group
Environment, National Physical Laboratory

ISSN 2059-6030

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

Approved on behalf of NPL by David Gibbs, Group Leader

CONTENTS

GLOSSARY

AOI	Area Of Interest
BIPM	Bureau International des Poids et Mesures
CEOS	Committee on Earth Observation Satellites
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
FAO	Food and Agriculture Organisation
FOREST	Fully Optimised and Reliable EmissionS Tool
FOTO	FOurier transform Textural Ordination
fPAR	fraction of Photosynthetically Absorbed Radiation
GCP	Ground Control Point
GPS	Global Positioning System
GUM	Guide to the expression of Uncertainty in Measurement
ISO	International Organization for Standardization
JCGM	Joint Committee for Guides in Metrology
LAI	Leaf Area Index
LiDAR	Light Detection And Ranging
LPV	Land Product Validation
MCA / MCS	Monte Carlo Analysis / Monte Carlo Simulation
MODIS	Moderate Resolution Imaging Spectroradiometer
MRV	Monitoring, Reporting and Verification
NDVI	Normalised Difference Vegetation Index
NMI	National Measurement Institute
NMO	National Measurement Office (UK)
PDF	Probability Density Function
QA	Quality Assurance
QC	Quality Control
RADAR	RAdio Detection And Ranging
REDD	Reducing Emissions from Deforestation and forest Degradation
REDD+	Reducing Emissions from Deforestation and forest Degradation, and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks, in developing countries.
RMSE	Root Mean Square Error
SAR	Synthetic Aperture RADAR
SI	International System of Units
TOA	Top-Of-Atmosphere
UAV	Unmanned Aerial Vehicle
WGCV	Working Group on Calibration and Validation

EXECUTIVE SUMMARY

There is a growing demand for trusted, fully integrated and cost effective Measurement, Reporting and Verification (MRV) services by forest carbon stakeholders. Forest monitoring projects and carbon initiatives require accurate and traceable data on forest parameters.

The Fully Optimised and Reliable Emissions Tool (FOREST) project aims to facilitate the delivery of information about forest condition and forest condition change through a suite of modelled and satellite-derived products (including both optical and RADAR). Following the Quality Assurance (QA) Framework, produced as part of this project, the forest products developed by the project partners will, for the first time, be provided with fully traceable detailed information concerning the processing steps and ancillary data input, as well as uncertainty characterisation and independent validation (where available). This will enable end-users the means to identify which data products are most suitable for their application.

As part of this process, traceability diagrams for each of the FOREST products have been developed (see http://projects.npl.co.uk/forest/FOREST.html). These have been used to identify the processes undertaken and the extent to which uncertainty has been considered and propagated through to the final product. Product evaluation and validation procedures have also been assessed. The aim of the traceability chain method of reporting is to identify areas in the processing algorithms of FOREST products where the traceability of the products may be improved, so in turn increase user confidence in the information.

Assessment of the processes, through the production of traceability chains for each of the FOREST products, has identified several recommendations where potential improvements could be made at this initial stage of product QA analysis.

1 INTRODUCTION

There are many methods available for deriving information about forests from satellite datasets. For users to be able to effectively determine if a data product is suitable to their needs, the data must be quality assured. This means that data products must be traceable (including full metadata and information concerning their production), and provided with quantifiable uncertainties. The generation of such quality-assured FOREST data products will enable comparability between different products and in situ data sets, as well as the assessment of products in relation to their intended applications. Forest products with embedded QA information would enhance end users' confidence (i.e. scientists and service customers) in the data for their intended application.

The derivation of full and robust uncertainty budgets for earth observation datasets is a relatively new concept in the field and many of the uncertainties associated with the data sets and processing routines have not yet been rigorously explored. The partners within the FOREST project have therefore, understandably, not been in a position to be able to provide such uncertainty budgets for their datasets. Therefore, this document will assess the methodologies (where appropriate) to identify key areas for future improvement in the product QA procedures. Recommendations identified in this report will begin to outline approaches to QA integration within the forest products and help build towards full uncertainty analysis. The assessment will consider all datasets and processes employed in their development, together with the evaluation of how uncertainty may be estimated and propagated through to the end product.

Traceability diagrams have been employed as a clear and robust approach to documenting the nature, provenance and relative rigour of the QA of the procedures used in the generation of forest product maps by the project partners. The importance of a traceability chain when attempting to estimate any form of uncertainty budget cannot be over emphasised as it allows the clear and systematic identification of potential uncertainties contributing to the overall budget. In an ideal case, a quality-assured product should include the identification, quantification and propagation of associated uncertainties for a specific dataset through the entire product procedure.

1.1 OBJECTIVE

The purpose of this recommendations document is to facilitate FOREST product developers in the generation of well documented products which provide end users with information on the quality of products, to aid decision-making on which FOREST product to use – i.e. establish which product is fit-for-purpose and the most suitable for their use.

1.2 SUMMARY OF RECOMMENDATIONS

A full uncertainty analysis of the FOREST products requires detailed information on each of the input datasets and processes carried out in their development. Uncertainties associated with datasets are often 'missing' due to a number of limitations, such as radiometric calibration uncertainties on the raw satellite data not being provided by the data distributor, and consequently there is a 'broken link' in the traceability chain. Consequently none of the project partners were able to implement a full uncertainty budget. However, a number of recommendations were identified from the existing methodologies undertaken, which should be considered by all partners in future in order to enable a full assessment of the quality of FOREST products.

Recommendation 1:	Clearly define end user requirements in terms of products and their accuracy.
Recommendation 2:	Document all processing steps to produce a full traceability chain including processing algorithms, ancillary data input and uncertainty characterisation.
Recommendation 3:	Where uncertainties have not yet been directly calculated, justify the use of the selected processing steps and datasets, compared to alternative approaches.
Recommendation 4:	Bias caused by subjective judgement should be considered in uncertainty analysis
Recommendation 5:	Sensitivity analyses should be conducted to identify the significance of different input parameters.
Recommendation 6:	Propagate uncertainties through the processing models using a Monte Carlo method (all inputs should have uncertainties assigned to them).
Recommendation 7:	Use error estimation to assess classification accuracy as it is difficult to estimate uncertainties for thematic datasets.
Recommendation 8:	Product developers should undertake comparison of products to identify differences in addition to a full validation using independent datasets.
Recommendation 9:	Validation should be undertaken using appropriate sampling strategies and traceable ground validation measurements.

2 RECOMMENDATIONS DISCUSSION

2.1 RECOMMENDATION 1: USER REQUIREMENTS

At the outset of any Monitoring, Reporting and Verification (MRV) programme, the requirements of the end user should be clearly identified as the entire product development process should be grounded by these needs. Requirements to be clearly identified should include for example:

- What parameter is required?
- At what spatial resolution should the parameter be provided?
- At what temporal resolution should the parameter be provided?
- What level of accuracy or uncertainty is acceptable in the parameter?

A clear definition of end user requirements should be used to determine the input datasets employed. For example, carbon certification schemes are increasingly becoming more stringent, such as requiring high accuracy classification to distinguish between land cover types. This requires higher spatial and temporal resolution satellite data for land use change mapping to comply with MRV requirements.

End user requirements should also determine whether local datasets (tailored to specific case studies) as opposed to default values, are appropriate. For example, the use of climatology datasets within models can limit the output unless local scale data can be acquired. When FOREST products are used by local-scale Reducing Emissions from Deforestation and forest Degradation (REDD) projects, then the input datasets used should be at an appropriate scale whenever possible, i.e. local data or high resolution. In the case where these input data sets are not available, justification for the use of existing data should be provided along with detailed information as to how the use of the selected dataset will affect the quality and uncertainty of the final product.

In some cases, the end users of a product will be well defined and therefore, these requirements can be easily captured. However, where end users are not predefined or may come from many different backgrounds, the requirements from these different areas should be sought and amalgamated in some way to drive the requirements of the product.

2.2 RECOMMENDATION 2: DOCUMENTATION

Before the uncertainty of forest products can be assessed, the full traceability chain through which it would be propagated, needs to be documented, as stated as the first stage in the QA Framework (see *Quality Assurance Framework for Forest Monitoring Programmes* - Barker et al, 2015). Figure 1 provides an example of the top level of the traceability diagram for the production of above-ground biomass using the ORICHIDEE-CAN model.

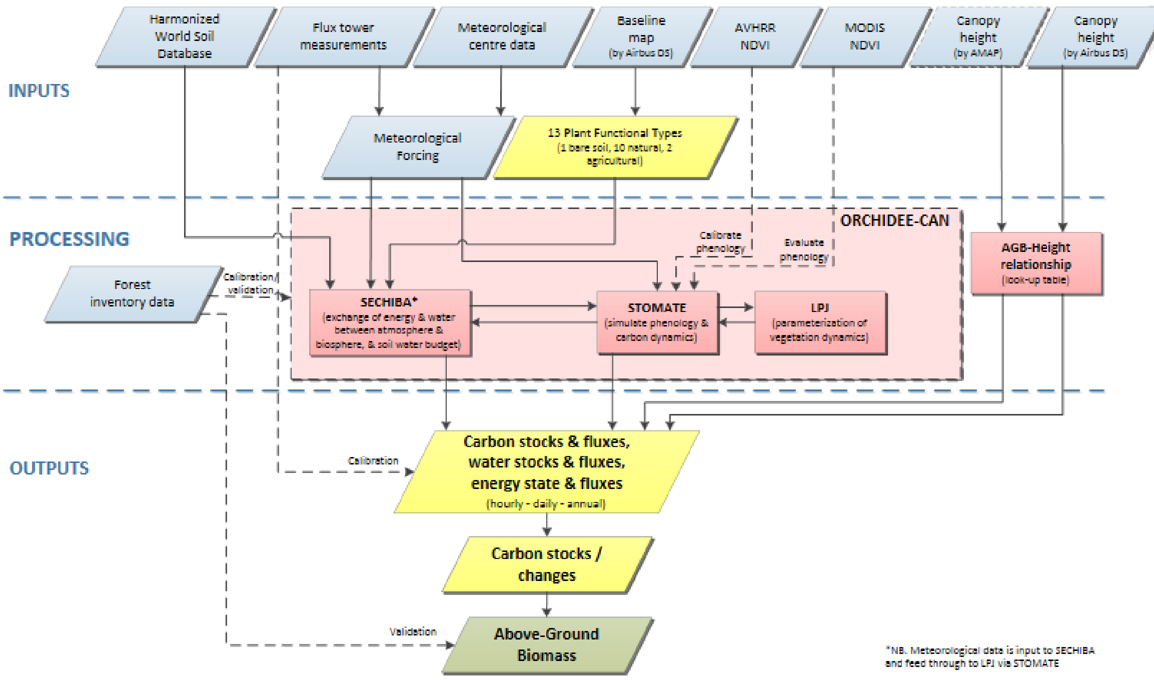

Figure 1 – Top-level traceability diagram summarising the procedures leading up to the generation of the above-ground biomass product by LSCE

It is key to ensure that all recorded metadata is current and accurate. All ancillary datasets, procedures and feedbacks should be fully documented (see Figure 1). This includes any assumptions, simplifications and approximations used in each processing technique, with a focus on those deemed to make a significant contribution to the overall processing scheme. An example identified within the land cover mapping product (by Airbus DS), is the physical assumption within the PROSPECT and SAIL model of *Lambertian* broad-flat leaves, horizontally homogenous plant canopies, or the assumption in LOWTRAN 7 model of a horizontally homogenous atmosphere and small elevation variation within a scene (van den Bosch, 2015). Such assumptions should be clearly documented so that end users are aware of the limitations of the methods utilised. This also allows data processors or end users to determine if the model can suitably interact with their applications and, where possible, identify model improvements or alternative methods which may overcome limitations in the model.

Information regarding the production of input datasets should also be recorded, where available, to ensure the uncertainties associated with these are propagated through to the final forest product. This is often overlooked as it is sometimes difficult to retrieve the required information from external organisations. For example, if the development and validation of the WorldDEM™ database (www.geo-airbusds.com/worlddem/) by Airbus Defence and Space (DS) was available, a more accurate assessment of its associated uncertainties would be possible.

Complete traceability chains for FOREST products are provided on the website (http://projects.npl.co.uk/forest/FOREST.html) as an example of the level of detail required in a traceability diagram to allow the evaluation and propagation of uncertainties through the processing chain.

2.3 RECOMMENDATION 3: JUSTIFICATION OF PROCESSES

Increasingly, end users require evidence in selecting the most appropriate approach for calculating forest products such as carbon stock or above-ground biomass for their product application, i.e. which method or data combination of these, whether modelling, RAdio Detection And Ranging (RADAR) or optical satellite. The inclusion of uncertainties can facilitate end users in this decision-making process by providing robust QA information. For example, whilst the Harmonized World Soil Database (www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/) (used as an input dataset to ORCHIDEE-CAN), may be considered the best available soil dataset, there is no classification uncertainty metric provided, thus justification of its use over and above an alternative becomes difficult. Therefore, where uncertainty assessments are not available to distinguish between similar products developed through different means, an explanation as to how decisions were made to use one dataset or process over other available options becomes the main source of information, upon which the user can base the justification of their use of the specific product.

Reasons for compromises in using one method over another should be explicitly stated. For example, in the case of the land cover mapping by Airbus DS, some discussion as to why the SAIL model was used and a 3D radiative transfer model not employed would benefit the user.

An example of justification within the FOREST project could be made for the implementation of the LOWTRAN 7 model. This atmospheric model could be justified through its ability to overcome a number of the limitations introduced by other atmospheric correction techniques, some of which have limited capabilities for vegetated scenes and consequently would not be appropriate for the assessment of forested land within the FOREST project. This atmospheric model does not require identification of spectrally neutral areas with no absorption features in their spectra (i.e. *a priori* knowledge of the scene), nor does it require ground measurements acquired at the same time as the image (van den Bosch, 2015).

2.4 RECOMMENDATION 4: BIAS

The effect of subjective judgements (i.e. the use of expert opinion) should also be considered in the quality assessment of products. Within the FOREST project, visually controlled and manually reshaped classification maps are generated. Where visual inspection of images or supervised classification methods are employed, for instance within the Overland tool or the Support Vector Machine by Airbus DS, respectively, there may be some bias between individuals assessment. Estimating traceable uncertainties introduced by human judgement are difficult to quantify and, therefore, these cannot be propagated through to the final product in this case. However, this should be quantified in terms of a clear assessment methodology and where open to diverse interpretation, intermittent cross-comparison exercises undertaken to normalise the bias of individuals, and so assess the spread in the interpretation.

Unsupervised classification such as that performed by the FOurier transform Textural Ordination (FOTO) method employed by AMAP, would benefit from local knowledge during its initial set up. Local knowledge of forest variation can be used to determine an appropriate analysis window size, however the introduction of human judgement and its potential bias would need to be considered.

2.5 RECOMMENDATION 5: SENSITIVITY ANALYSIS

Although uncertainties have not been fully propagated through the forest products in this project, a number of them have evolved with sensitivity and validation studies aimed at highlighting improvements that could be made and attempting to constrain sources of error.

Significant improvement in uncertainty analysis has been achieved through the use of statistical techniques to look at product's sensitivity to input datasets. The initial step in a sensitivity analysis being the identification of parameters the forest product / output is most sensitive to. For example, the Valade et al. (2013) study carried out a sensitivity analysis of ORCHIDEE using the Morris method (Morris, 1991) that was specific to looking for the most important parameters for simulating LAI, and using input parameters from an agronomical model for sugar cane. Parts of the sensitivity analyses could thus be adapted to the recommended uncertainty evaluation methodology detailed below (including identification of parameters for which a Monte Carlo simulation could be carried out).

Where optical satellite datasets are used, targeted sensitivity analysis can identify parameters which account for most of the variability in canopy reflectance, by keeping all parameters constant and varying one. This could be achieved through a Monte Carlo simulation. Jacquemoud (1993) varied the inputs to the PROSPECT/SAIL model one-by-one to assess each parameters impact on the canopy reflectance. However, it is important to note that simple parameter-by-parameter sensitivity studies do not mimic natural conditions as parameters are often independent, and the combined change in two parameters may produce a greater effect than the sum of the effects expected from either parameter alone.

2.6 RECOMMENDATION 6: UNCERTAINTY PROPAGATION

2.6.1 Introduction

The previous recommendations have clearly highlighted that to fully assess the suitability of a product for a particular application, evaluation of the uncertainties associated with the end product is useful. However, it is noted that the evaluation and propagation of uncertainties for earth observation products is a currently evolving discipline. Therefore, the information presented below is aimed at aiding in the development of an uncertainty budget for forest products and cannot be considered an exhaustive guide.

2.6.2 Identification of Uncertainties

To allow the full and metrologically robust propagation of uncertainties through the entire processing chain, both the input datasets and the processes undertaken need to be clearly understood. A potential aid to this process is the production of a traceability chain describing the inputs and processing undertaken to derive a product (such as that shown in Figure 1).

The traceability chain can then be utilised to identify, through careful consideration, the potential sources of uncertainty associated with the product. These may be from input parameters, for example those introduced by the WorldDEM Digital Surface Model (DSM), assumptions made within the processing, for example the assumptions relating to the atmospheric correction or process, or through the processes themselves, for example, spatial interpolation.

2.6.3 Estimating Uncertainties

Once the sources of uncertainty have been identified, efforts should be made to estimate the associated uncertainties. This may be achieved through consideration of the available data, reference materials, expert knowledge or other information. It is noted that it may not be possible to estimate the uncertainties associated with each of the identified sources of uncertainty; where this is the case, details of why an uncertainty cannot be estimated should be provided.

For evaluating the standard uncertainty for the outputs of simple models, i.e. a model for which a linear approximation is adequate, refer to the steps outlined in the Guide to the Expression of Uncertainty in Measurement (GUM) (JCGM, 2008a) to produce a combined standard uncertainty for forest products (see Figure 3 – equation to combine linear uncertainties).

Where complex models (perhaps containing non-linear steps) are utilised as part of a processing chain, for example in atmospheric correction codes, it may be less reliable to estimate the uncertainties associated with the output of the model through using the process outlined above. Therefore, it is necessary to consider an alternative method. One useful method of achieving this is Monte Carlo Analysis (MCA), which is allowed for in a supplement to the GUM (JCGM, 2008b).

MCA for uncertainty evaluation involves determining the Probability Density Function (PDF) associated with each of the input parameters then running simulations of the process(es) being studied, each time taking a random draw from the input PDF as the value of the parameter. A PDF can then be constructed from the outputs of the MCA and this can be used to determine the uncertainty associated with the output of the process. This can be represented as shown in Figure 2 below where the definitions are the same as for Figure 3, and the function *f* represents the model, which can be treated as a black box.

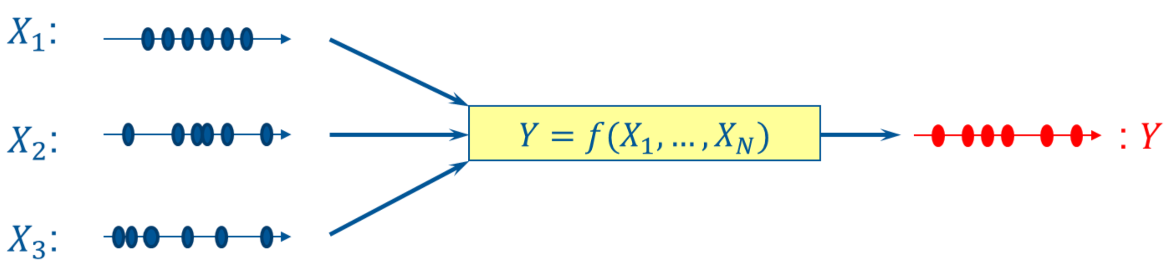

Figure 2 – Monte Carlo Analysis approach to uncertainty analysis.

2.6.4 Combining and Expanding Uncertainties

The combination of uncertainties can be achieved using the Law of Propagation of Uncertainties which can be found in the GUM as detailed in BIPM et al. (1995) and JCGM (2008a). This can be represented as shown in Figure 3 below where: x_1,x_2,x_3 are estimates of the input parameters X_1,X_2,X_3; $u(x_1)$,$u(x_2)$,$u(x_3)$ are the uncertainties associated with those estimates; v_1,v_2,v_3 are the degrees of freedom attached to each uncertainty; c_1,c_2,c_3 are the sensitivity coefficients; u(y) is the standard uncertainty associated with the estimate (y) of the output parameter (Y); and v_{eff} the degrees of freedom attached to u(y).

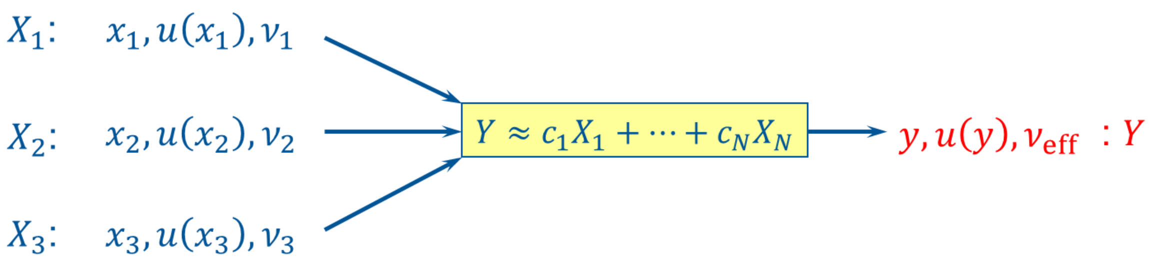

Figure 3 – Linear combination of uncertainties as provided by the Law of Propagation of Uncertainties.

The uncertainty estimate provided by the Law of Propagation of Uncertainties is the standard uncertainty; this should be expanded to the expanded uncertainty using the appropriate coverage factors.

2.6.5 Strategies for Minimising Uncertainties

Where possible, the most suitable input datasets should be selected. The uncertainty associated with models such as SAIL within the Overland tool can be reduced through the selection of forest biophysical parameters which can be directly measured in the field e.g. estimates of Leaf Area Index (LAI) or chlorophyll content, as opposed to the Normalised Difference Vegetation Indices (NDVI) which relies on indirect (inferred) measurements of the target quantity. Appropriate selection of parameters can allow a more accurate validation when comparing modelled and 'true' values.

2.6.6 Summary

There are two main recommendations for uncertainty propagation through complex models such as those within the FOREST project, including ORCHIDEE-CAN and LOWTRAN. Firstly, all inputs need to have an uncertainty assigned to them, and the issues of assigning uncertainties to classified map inputs have to be addressed. Secondly, when this prior recommendation has been followed, the use of standardised uncertainty evaluation techniques (including MCA) should be utilised to develop a full uncertainty estimate.

2.7 RECOMMENDATIONS 7 & 8: CLASSIFICATION ACCURACY & COMPARISONS

It is not currently practical to provide a comprehensive uncertainty budget for land cover and land cover change products for a number of reasons, including the lack of best practice guidance on how to quantify uncertainty associated with classification algorithms. Classification uncertainty (for supervised algorithms) is yet to be fully understood. Uncertainties of a discrete value i.e. thematic grouping such as a land cover type, cannot be quantified. However a classification accuracy study could consider the training dataset size in order to see if there was any significant improvement in the algorithm output with a larger training dataset (Olofsson et al, 2013; Olofsson et al, 2014).

If the acquisition dates of images from different satellite sensors is relatively similar i.e. images are acquired within a week of each other, then the accuracy of supervised classification can be compared between different satellite input datasets. However their approach to addressing issues arising from non-identical collection conditions, i.e. illumination and viewing angle geometries, changes in the atmosphere etc. should be considered as this can cause variation in results.

One method for assessing the classification is the comparison between different datasets which can identify where there are discrepancies. However this is not a measure of accuracy against true ground data, only an assessment of two different datasets.

2.7.1 Classification Error

Land cover change products should utilise methods suggested by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) Land Product Validation (LPV) for assessing error in data products where possible. Classification error acts as a first step towards an uncertainty budget for thematic data. A confusion matrix is commonly used to display the number of pixels considered and the number of pixels correctly or incorrectly classified. Statistical techniques are often employed for error estimation of end products. For example, the statistical relevance of the land cover/land cover change map by Airbus DS was assessed (calculation of the root mean square error), and the accuracy of the SVM supervised classification algorithm was also assessed.

2.8 RECOMMENDATION 8: VALIDATION

Validation is the process of assessing, by independent means, the quality of the data products derived from those system outputs. Validation of a satellite-derived dataset against field sites is typically used to assess the accuracy of the product that is extrapolated throughout the satellite product geographical extent. It is a necessary step to ensure a product is realistic by comparing it with ‘true’ values. Validation of datasets is imperative if their quality is to be known and the datasets are to be improved.

Validation of models (e.g. PROPSECT and SAIL) over forests is often limited, though several studies tend to be conducted over crops where in situ measurements of canopy structure can be obtained. For example, Andrieu et al (1997) and Danson and Aldakheel (2000) compared modelled and measured spectra for sugar beet crops. The coverage and number of measurements taken for validation datasets is inevitably limited. Where input information is different or incomplete, the comparisons cannot be considered full validations, but rather are used to identify trends in the data (Jacquemoud et al, 2009).

The ability of the SAIL model to estimate canopy biophysical parameters (chlorophyll content and LAI) was evaluated in Jacquemoud et al (2000) with regards to soybean and corn, by testing the model’s performance on real data. Reflectance simulated using the model and fitted parameters were compared to the corresponding in situ measured reflectance, measured by an LAI plant canopy analyser and a chlorophyll meter. Jacquemoud et al (2000) evaluated the fit using the root mean square error (RMSE) and accuracy as the mean distance from the measured values of the biophysical parameters to the global minimum. However it is key to remember that this approach does not take account of uncertainty nor propagate it. It merely provides an error assessment and degree of consistency of the final product.

For the estimation of uncertainty associated with satellite-derived products, it needs to be validated against an independent reference which contains information regarding its quality. However, validation is not always operationally feasible. For example, for land cover change mapping it is not always possible to acquire appropriate imagery e.g. TerraSAR-X before and after logging (Baldauf, 2013). However it is possible to target field trials in degradation hot spots for verification purposes.

Therefore, justification of process and which data has been validated and when, needs to be documented.

Within the FOREST project, validation datasets were not available in the case of mapping the canopy height using RADAR data. An alternative Digital Terrain Model (DTM) such as a LiDAR DTM was not available for validation of the WorldDEM DTM. The WorldDEM DTM and DSM used in the canopy height model by Airbus DS are the least accurate but most feasible datasets due to their global coverage. Currently there is an assumption of no or marginal penetration into the forest canopy as the WorldDEM does not account for this.

2.8.1 Sampling strategy

Validation datasets should be quality assured, and thus ground validation measurements require appropriate sampling strategies, as well as traceability and uncertainty estimation. Sampling strategies are often driven by particular constraints, for example, the resources available, as was the case for AMAP within the FOREST project. The comprehensiveness of the sampling strategy not only impacts the effectiveness of the validation process, but also affects the uncertainties associated with the validation process. For example, the uncertainty associated with the validation dataset may be reduced (if certain conditions are met) proportionally to the number of samples taken.

In addition, to prevent sampling strategies from producing biased results, an adequate number of spatially variable samples is needed. High density, stratified sampling of all study sites was not carried out within the FOREST project and has consequently resulted in the production of a limited and non-representative dataset. For example, not all forest types were sampled during in situ data collection and thus the datasets used in the supervised classification procedure carried out on the TanDEM-X dataset by Airbus DS was not considered representative. Areas of natural land cover change should also be considered as higher density sampling may be required in areas, for example, of natural erosion of coastal mangroves (Stach et al, 2015).

There is a need for an improved field sampling strategy via a random, stratified or systematic approach. The sampling strategy should consider the density of samples as well as spatial resolution. To improve the sampling strategy, prior information on the area of interest, as was available in the FOREST project, should be taken into consideration to better inform and thus design the sampling approach, i.e. knowledge of variation and use of statistical techniques to address the variation (Condit, 2008). Whilst sampling based on experience introduces bias, unlike a random sampling approach, it can lead to better informed selection of measurement locations to ensure a representative sample of each forest type within an area of interest (i.e. one can consider the heterogeneity of the forest). Sampling protocols for different measurands do exist, however standardised methodologies are not always appropriate due to the diversity of sites they attempt to cater to and the lack of local considerations.

Bias is introduced through the preferential selection of areas of forest with similar structure. In the case of the in situ dataset collected by AMAP within the FOREST project, this was old growth, in which a large number of samples had been collected.

2.8.2 Scaling

Upscaling in situ datasets at the leaf level to the level of satellite-derived products at the canopy scale is common in forest monitoring programmes (e.g. land cover products by Airbus DS). Satellite-based parameters are aggregated over a minimum area equal to the sensor's spatial resolution, whilst in situ validation measurements tend to be point measurements. Thus finer scale datasets are used in the validation of coarser scale products, and can rely on spatial homogeneity for upscaling. Sampling strategies need to be able to estimate the aggregated parameter values using an adequate number of spatially distributed measurements over the entire area of interest (AOI) (Widlowski, 2010). However at present, it is difficult to assign an uncertainty caused by errors in scaling.

2.8.3 Calibration

Assumptions within the algorithms related to the atmosphere and Earth's surface are common and difficult to overcome. These assumptions cause deviations between the values estimated by the product and the true value, and thus should be documented and included within any consideration of uncertainty.

Calibration of instruments (e.g. a chlorophyll meter) provides an uncertainty associated with the measurand. Parameters such as the fraction of photosynthetically absorbed radiation (fPAR) within the ORCHIDEE-CAN model could be validated using measurements taken by calibrated instruments. Calibration of instruments before and after use would enable us to understand their physical properties and deterioration, and thus understand the uncertainty associated with the measurement.

The training datasets of models such as LOWTRAN within the Overland tool need to be fully understood before the model's uncertainty can be properly assessed. Models should also be regularly updated via calibration using the most appropriate data available. For example, the phenology process described in the STOMATE model within ORCHIDEE-CAN could be updated using satellite data. Moderate Resolution Imaging Spectroradiometer (MODIS) data could be used for calibration to correct the growing season lengths (MacBean et al, 2015).

In summary, independent validation with good quality ground truth data is imperative and needs to be improved for the FOREST products. It is acknowledged that collecting this information is difficult due to constraints including but not limited to; time, cost, the trade-off for reduction in uncertainty and in obtaining high resolution images when needed. Consequently, a further recommendation would be aimed at funding bodies to ensure an adequate budget is secured for a validation component in the generation of products.

3 CONCLUSION

Full traceability is an extensive task for any MRV product. This recommendation document provides guidance on traceability and the consideration of uncertainty relating to procedures in the generation of forest products. Uncertainty propagation through an entire product traceability chain is difficult yet not unachievable. This document begins to identify the importance of documenting the steps where uncertainty is introduced and how some assessments of error and/or uncertainty are the first step towards development of a full uncertainty budget in future.

4 REFERENCES

Andrieu, B., Allirand, J.M., Jaggard, K.W. (1997). Ground cover and leaf area index of maize and sugar beet crops. *Agronomie,* Vol. 17, pp. 315-321.

Baldauf, T. (2013). Monitoring Reduced Emissions from Deforestation and Forest Degradation (REDD+): Capabilities of High- Resolution Active Remote Sensing. *Staats- und Universitätsbibliothek Hamburg: Germany.*

Barker, A., Origo, N. and Nightingale, N. (2015). *Quality Assurance Framework for Forest Monitoring Programmes.* NPL Report COM 1. ISSN 1754-2944.

BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML. (1995). *Guide to the Expression of Uncertainty in Measurement* (Geneva, Switzerland: International Organisation for Standardisation).

Condit, R. (2008). *Methods for estimating above-ground biomass of forest and replacement vegetation in the tropics,* Center for Tropical Forest Science Research Manual. Washington DC, USA.

Danson, F.M. and Aldakheel, Y.Y. (2000). Diurnal water stress in sugar beet: Spectral reflectance measurements and modelling. *Agronomie*, Vol. 20, pp. 31−39.

Jacquemoud, S. (1993) Inversion of the PROSPECT+SAIL canopy reflectance model from AVIRIS equivalent spectra: theoretical study. *Remote Sensing of Environment,* Vol. 44, pp. 281-292.

Jacquemoud, S., Bacour, C., Poilve, H. and Frangi, J.P. (2000). Comparison of four radiative transfer models to simulate plant canopies reflectance – Direct and inverse mode. *Remote Sensing of Environment,* Vol. 74, pp. 471-481.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P, François, C., and Ustin, S.L. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment, Vol. 113, pp. 56−66.

JCGM (2008a). Evaluation of measurement data – Guide to the expression of Uncertainty in Measurement (GUM 1995 with minor corrections).

JCGM (2008b). Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method.

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F-B., and Ciais, P. (2015). Using satellite data to improve the leaf phenology of a global Terrestrial Biosphere Model. *Biogeosciences Discuss*., Vol. 12, pp. 13311–13373.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, pp. 161-174.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E. and Wulder, M.A. (2014). Good Practices for Assessing Accuracy and Estimating Area of Land Change. *Remote Sensing of Environment*. Vol. 148, pp. 42-57.

Olofsson, P., Foody, G.M. Stehman, S.V. and Woodcock, C.E. (2013). Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, Vol. 129, pp. 122-131.

Stach, N., Salvado, A., Petit, M., Ingrassia, F., Joubert, P. and Deshayes, M. [Accessed 10/2015]

Monitoring land use and land use changes in French Guiana by optical remote sensing. Inventaire Forestier National (INF); Bron, France. Available from: www.researchgate.net/publication/248978036_Land_use_monitoring_by_remote_sensing_in_tropical_forest_areas_in_support_of_the_Kyoto_Protocol_The_case_of_French_Guiana

Valade, A., Vuichard, N., Ciais, P., Ruget, F., Viovy, N., Gabrielle, B., Huth, N. and Martiné, J-F. (2013). ORCHIDEE-STICS, a process-based model of sugarcane biomass production: calibration of model parameters governing phenology. *Global Change Biology: Bioenergy* 6(5), pp 606-620

Van den Bosch, J. [Accessed 08/2015]. Application of LOWTRAN 7 as an atmospheric correction to airborne visible/infrared imaging spectrometer (AVRIS) data. Jet Propulsion Laboratory/California Institute of Technology; California, USA. Available from: http://aviris.jpl.nasa.gov/proceedings/workshops/90_docs/7.PDF

Widlowski, J.-L. (2010). On the bias of instantaneous FAPAR estimates in open-canopy forests. *Agricultural and Forest Meteorology,* Vol. 150, pp. 1501-1522.