INTERNATIONAL CLIMATE INITIATIVE

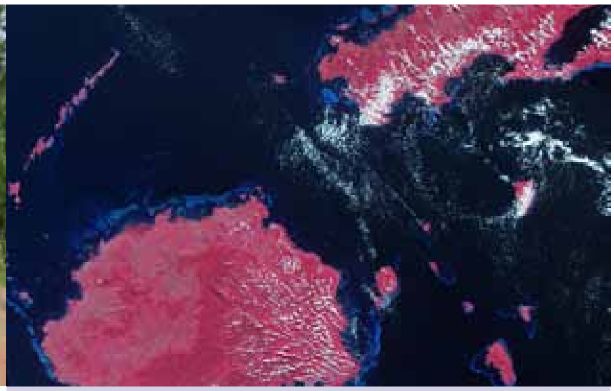
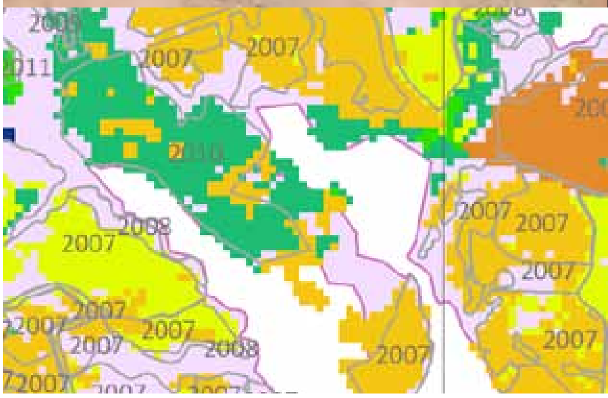
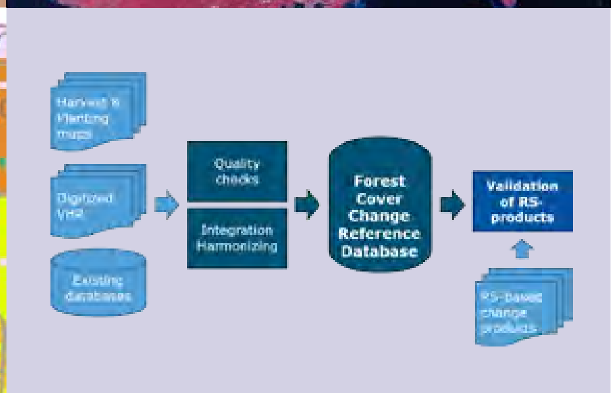
Regional project

Climate Protection through Forest Conservation in Pacific Island Countries

On behalf of:

Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety

of the Federal Republic of Germany

Reference Database and Validation Concept for Remote Sensing based Forest Cover Change Products in Fiji

MINISTRY OF FISHERIES & FORESTS
"Our Forest, Our Ocean, Our Future"

SPC Secretariat of the Pacific Community
Applied Geoscience and Technology Division (SOPAC)

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Reference Database and Validation Concept for Remote Sensing based Forest Cover Change Products in Fiji

March 2014

SPC
Secretariat of the Pacific Community
Applied Geoscience and Technology Division (SOPAC)

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

On behalf of:
Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety
of the Federal Republic of Germany

Prepared by: Johannes Eberenz (johannes.eberenz@gmail.com)
Johannes Reiche (johannes.reiche@wur.nl)

On behalf of: SPC/GIZ Regional Project
'Climate Protection through Forest Conservation in Pacific Island Countries'

P.O. Box 14041, SUVA, Fiji

Email: karl-peter.kirsch-jung@giz.de

Summary

Validation of Remote-Sensing-based forest-cover change products is crucial for obtaining uncertainties of the change monitoring. This is important for REDD+ monitoring and reporting. This study comprised an inventory of suitable reference data, design of a database and validation concept and a demonstrative implementation using the available spatial data infrastructure.

The inventory shows that for plantations and controlled harvest suitable reference data of good spatial and information quality is available, while data is missing on degradation and uncontrolled logging of indigenous forest. It is therefore recommended to set up additional permanent sampling plots, using clustered design stratified by forest type, change potential and accessibility.

Following the inventory and a literature review, a reference database was designed based on change events and baseline data. It allows extraction of reference data by spatio-temporal queries and can handle data derived from point and area surveys and from satellite imagery. The used forest cover change type definitions are based on the definitions currently used by the forestry department and adapted to the available data. It is recommended that Fiji develops more precise definitions for forest degradation and improvements. The validation concept can handle different sampling designs and bi-temporal, yearly or sub-yearly change information. Reported accuracy measures are the confusion matrix, overall accuracy, per class user's and producer's accuracy and the time-lag of change detection. Further recommendation regard better data management for transparency and reproducibility and building up of technical capacity sin the implementing agencies.

The study for a concept for a reference database and the validation for RS-based forest cover change products in Fiji was funded through the project Climate Protection through Forest Conservation in Pacific Island Countries of the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) on behalf of the German Federal Environment Ministry and performed in cooperation with Fiji Forestry Department Management Services Division (MSD) and the Applied Geoscience and Technology Division of the Secretariat of the Pacific Community (SOPAC) in Suva, Fiji.

Data Ownership

Datasets used in this research were provided by Fiji Pine Limited (FPL), Fiji Hardwood Cooperation Limited (FHCL), Fiji Forestry Department MSD and SOPAC as noted below. The use of the datasets is restricted to this research. Data ownership remains with the respective providing organizations of the datasets. Before the implementation of a forest change database, the terms and conditions of integrating the different datasets will have to be agreed upon by the owners of the datasets and the implementing organization.

Contents

List of Figures

List of Tables

Acronyms & Abbreviations

ALOS	Advanced Land Observing Satellite
BFAST	Breaks For Additive Season and Trend
CI	confidence interval
DB	database
DBDLC	database development life cycle
E/R	entity / relationship
ETM+	Landsat Enhanced Thematic Mapper
FAO	Food and Agriculture Organization of the United Nations
FCC	FCC
FHCL	Fiji Hardwood Cooperation Limited
FID	feature identification number
FPL	Fiji Pine Limited
GIS	geographic information system
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
IPCC	Intergovernmental Panel on Climate Change
ISO	International Standard Organization
ISOTC211	International Standard Organization technical committee for digital geospatial data
MRV	monitoring, reporting and verification
MSD	Fiji Forestry Department Management and Survey Division
NDVI	normalized difference vegetation index
OA	Overall accuracy
OGC	Open Geospatial Consortium
PSP	Permanent Sampling Plots
REDD+	Reducing Emissions from Deforestation and Degradation, enhancement of carbon stocks, conservation of forest and sustainable forest management
RS	remote sensing
SDI	spatial data infrastructure
SLC	Landsat scan line corrector
SOPAC	Applied Geoscience and Technology Division of the Secretariat of the Pacific Community
TM	Landsat Thematic Mapper
UNFCCC	United Nations Framework Convention on Climate Change
VHR	very high resolution
WFS	Web Feature Services

1 Introduction

1.1 Context and Background

Forest cover change (FCC) affects many ecosystem services such as carbon storage, biodiversity, water retention and climate regulation. Particularly, deforestation and forest degradation are major contributors to greenhouse gas emissions (see e.g. Houghton 2012; Foley et al. 2005). With the United Nations Framework Convention on Climate Change (UNFCCC) 2013 summit in Warsaw, the implementation of Reducing Emissions from Deforestation and Degradation, enhancement of carbon stocks, conservation of forest and sustainable forest management (REDD+) has been confirmed. This mechanism requires participating countries to closely monitor and report the state and change of their forest resources, including forest area change (UNFCCC 2013). Development and improving the methodologies for this monitoring effort is one major challenge for REDD+ participants (Asner 2011; Romijn et al. 2012)

One method to assess national forest carbon stocks is the so called stratify multiply (SM) approach, where forest is stratified by and each strata is assigned an average carbon density (Goetz et al. 2009). To derive carbon dynamics, national-scale forest cover change information is required, also called activity data (IPCC 2006). Remote Sensing (RS) offers great potential for providing forest cover change information (De Sy et al. 2012). Apart from its application in a REDD+ framework, forest cover change monitoring is also valuable for sustainable forest management in a broader sense.

In order to assess the accuracy of RS-based forest cover change information an validation (also called accuracy assessment) using reference data is necessary (see e.g. Foody 2002; Stehman and Czaplewski 1998). Apart from being crucial for method development, in the REDD+ context validation can be used to derive uncertainties of the estimated carbon dynamics. Reporting uncertainties is recommended by the IPPC in the higher 2 and 3 Tiers approaches (Maniatis and Mollicone 2010). However, the important issue of validation is often not treated with all the attention it requires, as literature reviews of major remote sensing journal show (Foody 2002; Olofsson et al. 2013). Validation of RS based change products is more challenging than of land cover classification production and approaches are less established (Foody 2002).

Fiji comprises a tropical archipelago located in the tropics in the South Pacific. About 53 percent of its land mass is covered by forests (Leslie and Tuinivanua 2010). Fiji is currently in the process of preparing for REDD+-readiness (FCPF, Readiness Preparation Proposal for Fiji, 2014, see Fiji-R-PP; Fiji Department of Forestry 2011) . This has triggered a range of monitoring and inventory efforts (as e.g. the ReCover project, see http://www.vtt.fi/sites/recover/). This report is the outcome of the collaboration between the University of Wageningen and the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) as part of the project Climate Protection through Forest Conservation in Pacific Island Countries. The project contributes to Fiji's efforts towards setting up a national REDD+ program. The GIZ cooperates with the Fiji Forestry Department (FFD) and the Secretariat of the Pacific Community (SPC). GIS related tasks are conducted mostly by the FFD Management Services Division (MSD) the Applied Geoscience and Technology Division of the SPC (SOPAC). The internship was conducted in liaison with those divisions. The total period of the work was October 2013 to March 2014, of which November – February where conducted on Fiji. The work supports the REDD+ national programme by assessing the possibilities for validating RS-based forest cover change products for Fiji.

1.2 Problem Definition

For Fiji, different sets of forest change data are available, ranging from inventory data, data on harvest, replantation and selective logging to very high resolution (VHR) satellite based imagery. However the different data sets lack a common format and are of unknown quality.

Reference data from different sources and in different data types need a common format and interface in order to be used for accuracy assessments. Reference data is often assumed to be of "gold standard" quality, neglecting possible spatial and thematic errors, which can bias the validation outcome (Foody 2002). It should therefore be thoroughly checked and steps should be taken to account for potential errors. Storing the data in a common format simplifies the validation.

Given harmonized reference data of known quality, the next challenge is the validation concept. The selection of suitable accuracy measures depends on the sampling design and the requirements of the users (Foody 2002). For Fiji, no validation concept fitting the data and the problem of forest cover change has been developed so far.

1.3 Study Objectives and Questions

This study aims to demonstrate the setup and use of a reference database for validating RS-based forest cover change products for Fiji. The final goal is a complete workflow covering the steps from data inventory to accuracy validation. The objectives of the research are thus:

1. Inventory of availability forest cover change data and assessment of its suitability for validating RS-derived change products.
2. Design of reference database that provides the framework to harmonize input different forest cover change datasets to assess forest change.
3. A concept to use this database for assessing RS-derived forest covers change, for both bi-temporal and yearly changes.

2 Study Site and Material

2.1 Study Site

2.1.1 Fiji

Fiji is located in the South Pacific Ocean and comprises more than three hundred islands. The biggest island Viti Levu with ca. 10 000 km² makes up more than half of Fiji's land mass, followed by Vanua Levu with ca. 5500 km² (Encyclopedia Britannica 2014). These two biggest islands are referred to as the main islands.

About 53 percent of Fiji's land is covered by forests, including indigenous tropical and dry land forest as well as hard- and softwood plantation. While the westward sides of the main island and the highlands receive high precipitation (2500 – 3800 mm/year), the leeward sides are dryer (1500 – 2000 mm/year). This leads to a contrast in vegetation type, with rainforest in the west and grassland and dry forest in the East of the main islands (Leslie and Tuinivanua 2010).

Most of the indigenous forest is owned mataqalis, traditional Fijian communal units. While logging in these areas is regulated by the forest department, subsistence and commercial agriculture also lead to deforestation and forest degradation. About 36 percent of the indigenous forest is preserved or protected but this doesn't exclude all deforestation (Leslie and Tuinivanua 2010).

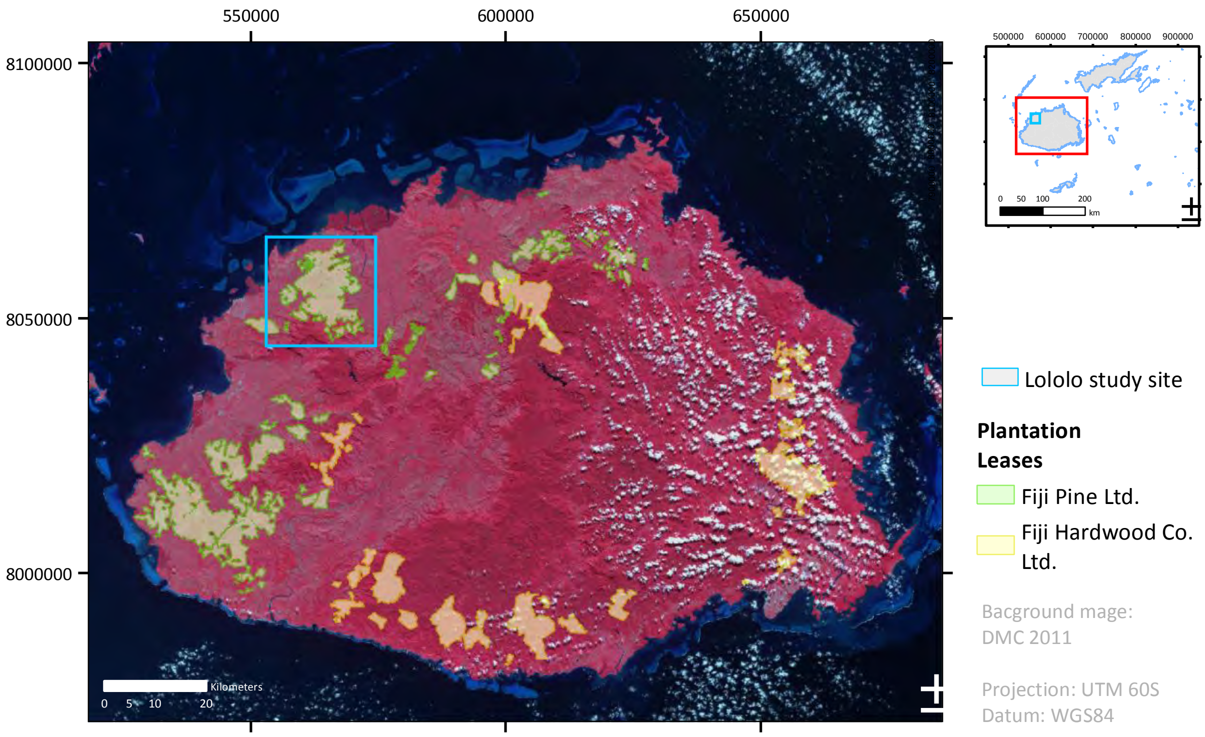


Figure 1: False color DMC mosaic of Viti Levu with plantation lease areas and Lololo study site extent.

State owned cooperation lease the timber plantation land from communal landowners. Softwood plantations (*Pinus caribaea*) make up ca. 2.5 percent of Fiji's total area. They are located on the leeward areas of the main islands and managed by Fiji Pine Limited (FPL). The Hardwood plantations are dominated by mahogany (*Swietania macrophylla*) and cover ca. 2.9 percent of Fiji's land mass. They are established on former rainforest areas in the eastern and highland parts of the main islands and managed by Fiji Hardwood Cooperation limited (FHCL) (Payton 2012; Leslie and Tuinivanua 2010). Figure 1 depicts the plantation lease areas on Viti Levu. a false color DMC satellite image of the two main islands of Fiji.

2.1.2 Lololo Study Site

The *Lololo* pine plantation lease area is located in the North-West of Viti Levu (see Figure 1) and managed by FPL. It covers ca. 16990 ha, of which ca. 60 percent is planted with *pinus caribaea* (FPL). Harvesting cycles are on average of 15 to 20 years long. Forest stands are completely harvested and then fully replanted. The area receives 1600 – 1800 mm precipitation per year (Leslie and Tuinivanua 2010). During the dry season from April to October, drought spells reduce vegetation growth. Major droughts occurred in 1998, 2003 and 2010 (Rina 2010). During the wet season, cyclones and heavy rainfall can damage the vegetation. In 2012 an extreme rainfall event caused loss of ca. 200 ha of pine plantation through landslides (Chaudhary 2012).

2.1.3 REDD Program

Since 2009, Fiji has a national REDD+ program. Together with international partners, the Fiji Ministry of Fisheries and Forests is aiming to establish a monitoring, reporting and verification (MRV) system and REDD+ activities. Payton (2012) recommends for Fiji to take up a SM approach with spatial explicit land cover conversion monitoring. This makes forest cover change monitoring one of the focus point of Fiji's REDD+ program.

2.2 Forest Cover Change Data

Raising an inventory of the available forest change data is an essential step of this research (see *Methods – Forest Cover Change Data Inventory and Quality* Control). The data includes:

- Indigenous forest harvest data
- Permanent sampling plots
- Plantation (hardwood and pine) harvest and planting data
- Change maps derived from RS imagery, including very high resolution data

Section *Results – Forest Change Data Inventory and Quality Control* gives the results of this inventory with more details on the separate datasets.

2.3 Remote Sensed Imagery

2.3.1 Landsat Data

Pre-processed multi-temporal Landsat Thematic Mapper (TM) 4 and 5) and Enhanced thematic Mapper 7 (ETM+) scenes with 30 m spatial resolution were provided by Johannes Reiche (Wageningen Univeristy). Pre-processing included cloud masking with FMASK (Zhu and Woodcock 2012) and atmospheric correction using LEDAPS (Masek et al. 2012). The scenes of path/row 075/72 and 074/72 cover most of Viti Levu (see Figur 2). For the years 1999 to 2012 ca. 150 images per scene where available (see Table 1). There is a data gap from 1993 to 1999 due to the failed Landsat 6 mission. The images show a great amount of cloud cover especially on the western side of Viti Levu, resulting in an average 77% missing cell values. From 2003 on, the slant line corrector failure (SLC) of Landsat 7 further increases the number of missing values. Figure 2 shows the number of valid observation per land pixel.

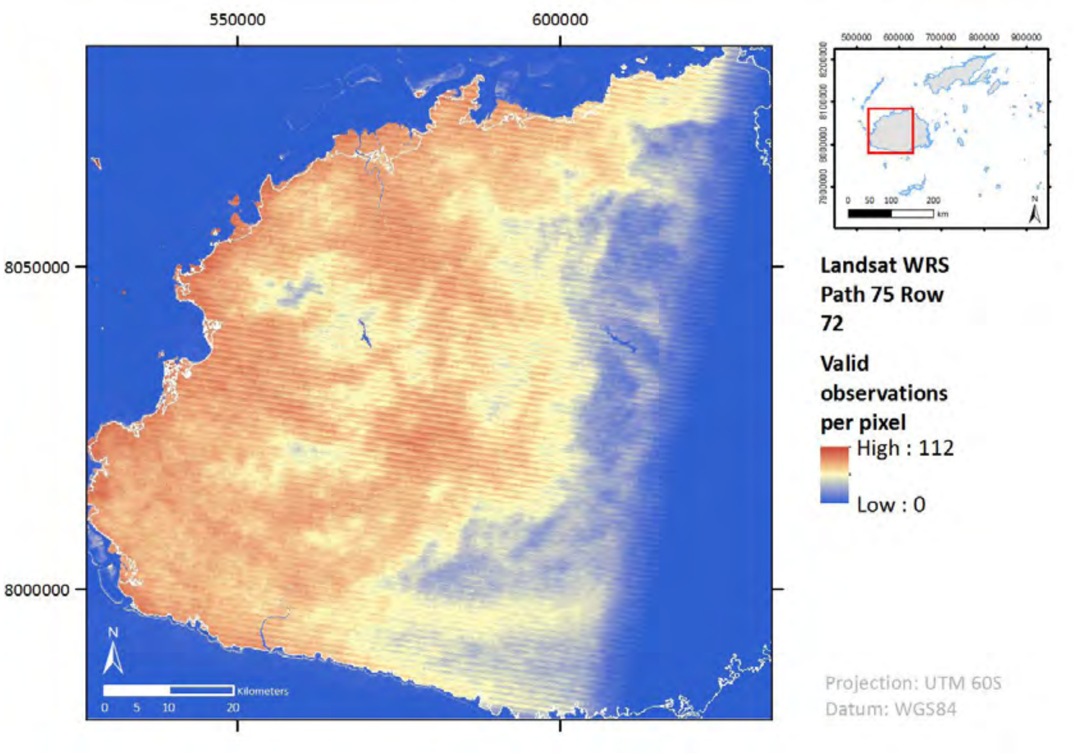


Figure 2: Valid observations per non-water pixel for Landsat WRS Path 75 Row 72, 1990-2012

Table 1: Landsat imagery: number of scene per sensors

Path/Row	Sensor	No. Of Scenes	Dates
075 / 72	TM	14	1990 - 1993
	ETM+ SLC-on	21	1999 - 2002
	ETM+ SLC-on	116	2003 - 2012

2.3.2 High and Very High Resolution Imagery

SOPAC granted access to a range of Quickbird and Wordlview 2 of the years 2003, 2004, 2010 and 2012 and 2013 respectively. Further Quickbird and Worldview images were accessed through GoogleEarth™. Where possible, the geo-registration of the VHR imagery was visually assessed using GIS layers and found to be of high quality.

The VHR imagery was used for visual interpretation and quality control only and is therefore not further described.

2.4 Remote Sensing Based Change Products

To demonstrate the use of the database, a RS-based forest cover change product was provided that was generated by Johannes Reiche (Wageningen University). It is based on Landsat normalized difference vegetation index (NDVI) time series data of 2006 to 2012 and covers the Lololo study area. The change product was derived using sing a pixel based time series analysis method (BFASTmonitor, Verbesselt, Zeileis, and Herold 2012). Pixels with an abrupt decline of forest cover were detected per pixel on a yearly base.

2.5 Additional Data

Additionally to the mentioned datasets, geo-information data such as land use, island coastlines, river and road networks and administrative boundaries where provided by the local partners.

3 Methods

3.1 Forest Cover Change Data Inventory and Quality Control

For the change data inventory, initially a working definition for forest cover change data has to be found. Based on this, potential FCC dataset can be selected.

The working definition of forest cover change is based on the definitions used by the local agencies.

The Fiji Forestry Department currently uses a forest definition adopted from the United Nations Food and Agriculture Organization (FAO). Forest is defined as *"Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ"* (Fiji Department of Forestry 2011; citing FAO 2004). This includes the plantations.

Forest cover change can take a variety of forms: harvest, deforestation and degradation, and replantation, afforestation, reforestation and regeneration. Following the FAO definitions, forest is considered as forest throughout the whole harvest cycle, if it includes replanting or regeneration. Harvest and replantation are a temporal change of land cover but not of land use, while deforestation and reforestation are a change of land use. Deforestation can have different drivers, such as expansion of agriculture or natural disasters, where the forest can't recover (FAO 2010).

The used FAO definition of forest degradation is less clear: "*Changes within the forest which negatively affect the structure or function of the stand or site, and thereby lower the capacity to supply products and/or services"* (FAO 2004; for a discussion see Simula 2009; Herold and Skutsch 2011). The FFD currently doesn't specify this further. Therefore, it's unclear if e.g. selective logging the indigenous forest is considered as forest degradation.

The available forest change datasets are explored and described. This includes spatial and thematic information and metadata. The inventory further includes searching for potential sources of additional data. Using RS-imagery, the data has to be checked for spatial and information consistency. This means checking the extend change type and time of the change areas by comparing with available VHR imagery using visual interpretation. Metadata is checked using auxiliary information. Quality problems in the dataset are stored explicit using quality flag attributes. Where quality control is impossible due to lacking data, this is also indicated.

All GIS dataset were original projected in Fiji 1986 / Fiji Grid[1], and are re-projected into WGS 84 / UTM zone 60 south[2] using the ArcMap project tool.

3.2 Database Design

A spatial database can be described as "a set of data describing the semantic and spatial properties of real world phenomena (temporal properties are also possible)" (Longley et al. 2005, chap. 29).

A full database development life cycle (DBDLC) contains following steps (Yeung and Hall 2007, chap. 3):

1. Initial Database Study
2. Database Design
3. Implementation and data load
4. Testing and evaluation
5. User education and training
6. Maintenance and monitoring

The scope of this research covers only steps 1 and 2. The demo implementation partially covers steps 3-5, however it really is only a prototype and thus should be seen as part of the design

The initial database study serves to define the objective and specify the data requirements. The points to cover are: Functionality, necessary Data, application interface and performance (Yeung and Hall 2007, chap. 8).

The database design follows a two level approach (Longley et al. 2005, chap. 29): The first step is an analysis or conceptual data model. It describes entities, relationships and integrity rules. Special attention is paid to the representation of temporal information (Longley et al. 2005, chap. 8). The model is represented in an entity / relationship (E/R) diagram.

The second step is a demonstrative implementation. Here the technological environment is taken into account, resulting in a physical data model. Tables and fields are specified and software choices are made (West 2011). The demo implementation is then used to show how an example change product can be validated.

3.3 Data Integration and Harmonization

For the different input datasets, methods are developed for converting the data into the format of the database (harmonization) and for updating the database with the harmonized data (integration). The input datasets can be in both vector or raster format, and vary in thematic content, spatial and temporal support, resolution and extent and in information quality. These differences have to be acknowledged and will lead to a differentiated methodology.

[1] For a detailed description see **http://www.spatialreference.org/ref/epsg/3460/**

[2] See **http://www.spatialreference.org/ref/epsg/32760/**

3.4 Validation Concept

Taking into account the special aspects of validating change and the properties of the available reference data, the validation concept covers reference data sampling design and sample size, validation methods and proposes a set of accuracy measures (see e.g. Congalton 1991; Foody 2002; Stehman 2009).

Validating change data is complicated by two factors: the need to cover both space and time and the typically relative small extend of change areas (Stehman 2009). For a statistical sound validation, every location on the map to be validated must have non-zero probability of being sampled (probability sampling). In this way the validation represents the entire mapped area (Stehman and Czaplewski 1998). Using non-probability sampled reference data can causes bias. In this situation it is preferable to constrain the validation to subareas where probability sampling is possible (Stehman 2009).

The conditions of probability sampling are fulfilled either by

a. *Complete survey* or *wall-to-wall* sampling: reference data is available for the full extent of the product to be validated. The validation then becomes a map comparison, eliminating sampling errors and the need for stratified estimators. (Stehman 2009).

b. *Point sampling* following a random or systematic design which can be clustered and or stratified.

For point sampling, two additional problems arise: choosing the sample size and sampling error. Statistical calculations of an adequate sample size are not trivial (Congalton 1991). For large population size as it typically occurs with remote sensing, Stehman (2001) shows that 100 samples are sufficient to reach a small standard error of 0.05 for random sampling. For stratified sampling, Congalton (1991) suggest at least 50 samples per stratum as a rule of thumb. For large number of classes, this number should be increased to 75 to 100.

Useful information can also be derived if probability sampling is not possible, however in this case no statistical valid measure for the accuracy of the entire mapped area can be derived (Stehman and Czaplewski 1998; Stehman 2009). For the purpose of developing a change monitoring methodology rather than map validation, non-probability sampled reference data can be used.

4 Results

4.1 Forest Change Data Inventory and Quality Control

All available FCC datasets where explored and visually assed using VHR imagery in order to assess their suitability for use as reference data.

The data was generally stored locally at the different partners in files and local databases. Often there was no regular update and change tracking and no controlled backups.

Table 2 gives an overview of this inventory. For the Fiji Pine harvesting data a more detailed report is provided, as it is used in the demo validation.

Table 2: FCC datasets inventory overview.
Orange text indicates shortcomings of the datasets for the purpose of FCC validation. Until now, no change data derived from digitized VHR imagery exists.

Dataset:	Permanent sampling plots	Indigenous forest harvesting GIS data	Pine / Hardwood Plantation GIS Data	(Potential Digitized VHR imagery)
Source:	FFD MSD	FFD MSD	FPL, FHCL	Presently none
Temporal resolution (sampling interval):	Medium (2 years)	Fine (1 year)	Very fine (1 year or sub year)	Coarse (approx. 5 years)
Temporal coverage:	2010– 2013, continued	2005 – 2012, continued	Ca. 2004 – 2012, continued	Not yet available (2013 – ...?)
Temporal info. quality	High	High	High	High
Spatial resolution / sampling density:	Low (100 plots for main islands)	High (GPS data)	High (GPS data)	Very high (pixels < 10 x 10 m)
Spatial coverage:	Sample of main islands	Harvested areas of indigenous forest of main islands, but incomplete	Plantation area	Potentially complete main islands, some cloud cover
Spatial information quality:	Location accuracy medium (GPS errors)	Unknown, heterogeneous plots, GPS errors	Medium (offsets, heterogeneous plots)	Very high
Thematically quality:	High	Unknown	Medium (some years missing)	High, (depending on digitizing method)
Sampling design:	Systematic (regular grids)	Based on management activity	Based on management activity	Complete survey (apart from clouds)
Probability sampling	Yes	No	(No)	Yes

4.1.1 Fiji Pine Limited data

a) Landslides:

The dataset contains areas affected by landslides in the Lololo forest area. The data covers the year 2012 when a major rainfall event caused a large amount of landslides in the area (Chaudhary 2012). The spatial offset form visual comparison with LS TM imagery lies between 0—2 Landsat pixels (0—60 m). The area covered is quite large (ca. 200 ha of planted pine), hinting that not mapped landslides in other areas could bias validation results significantly.

b) Logged Areas

This dataset depicted harvested forest stands from 2004 to 2012 in the softwood plantations managed by FPL, which is also the source of the data. For Viti Levu, four major lease areas are covered (depicted by the "Forest" attribute). Spatial information was gathered by field surveys using satellite positioning. The dataset has 1170 stand entries covering ca. 9.570 ha. Each polygon has a stand ID, some of which are not unique. The temporal resolution is fine: quarterly logging periods are reported. For the Loloo lease there is no data before 2006.

The FPL logging was compared to the available RS imagery. Generally, thematic and spatial information are of high quality. For the Lololo area, more detailed control was performed: all plots harvested between 2006 and 2012 were checked against VHR imagery. Only very few areas without logging information show clear logging traces (see e.g. the yellow outlined plot in Figure 3). Some stands were apparently only partially harvested, resulting in mixed forest cover.

There are large area with reduced vegetation in 2010, possibly due to drought (see paragraph 2.2.1 Lololo Study Site and Rina 2010). These could affect change detection for this year.

The spatial information is also satisfying; the offset between the VHR imagery and the plot outlines was in a range of 0 – 30 m. This could be due to errors of the digitized map material or satellite positioning system. As described above, the geo-registration of the VHR imagery seems of high quality, contributing little to the perceived error.

Additionally to the quality control, the dataset is visually compared with the Landsat time series stack. This is not strictly speaking quality control, since the change products to be validated are derived from the Landsat data validated. However, it can help to understand the results of the validation.

The thematic information between the RS and the reference data is highly corresponding. 99 percent of the logging years and 93,7 percent of the logging period (quarter of the year) in the reference data are visible in the in the Landsat time series stack (See Table 3).

Table 3: Results of visual assessment of temporal information in FPL harvest data 2006 - 2012

Type of temporal information in dataset	Valid entries (plots)	No. of plots where dataset period overlap with logging period derived from visual interpretation
Logging Quarter of year	206	193 (93.7%)
Logging Year	208	207 (99.5%)

The spatial correspondence with LS data was found to be of varying quality: A spatial offset of 1 – 3 pixels (30 – 90 m) in varying directions is common (see Figure 3). These spatial offsets influence the validation results.

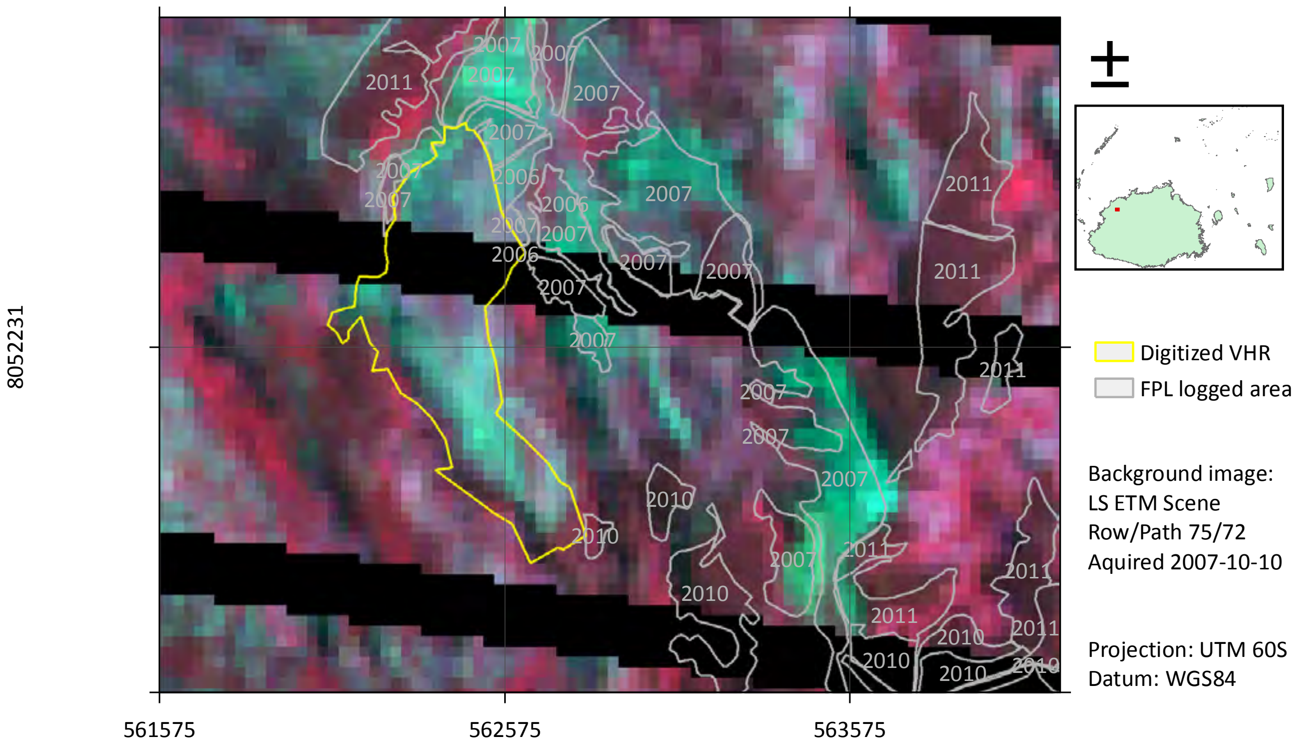


Figure 3: FPL logging data vs. Landsat imagery

c) Planting

This dataset depicts replantation of softwood Plantations by FPL. Coverage and attributes are comparable to the FPL harvest data. Polygons typically don't overlap exactly with corresponding harvested areas, indicating separate digitization. Available temporal information is mostly only the planting year, with more 70% than missing entries for the quarter attribute. The datasets covers a larger area (ca 35765 ha) and has more entries (1750) than the harvest data, mostly because earlier years are covered. The years that are missing in the harvest data for the Lololo study site area are covered. The years 2008 and 2009 have very few entries, while 2012 shows usual high number of planted areas. This can be explained by the economic difficulties the company faced (Lal 2013). Quality control using VHR imagery showed similar results as with the FPL logging data.

Comparison with the harvest data shows that in 2005 and 2006 most plots were replanted one to two years after harvest. Therefore it is possible to use the planted areas where harvesting data is missing (e.g. years 2006 and 2005 in the Lololo lease area).

4.1.2 Fiji Hardwood Ltd.: Logging Data

This datasets depicts harvested areas in the Fiji hardwood plantations for 2003 to 2013. The dataset contains 593 entries with 15 invalid entries. Comparing the data with VHR imagery shows high spatial and thematic quality. Some Plots show complete harvest, others selective harvest with skid roads and logged patches clearly visible. Harvest is visible in the spectral information still after several years. In the Landsat imagery, very little of the logging is visible. Also the hardwood areas are more affected by cloud cover than the pine plantations.

For both pine and hardwood, the plantation lease boundaries are given in the land use data. However, not the complete leased areas are currently stocked. (2012, chap. 4.2.3) therefore recommends deriving land cover based extent information for the plantations.

4.1.3 Indigenous Forest Harvesting Data

Harvesting data for the indigenous forest of the main islands is provided by the FFD MSD. The data contains of separate shapefiles per year (2003-2013). Between the files, the attributes are not handled consistently. Each plot can be uniquely identified by the logging registration number. For some plots also the removed volume and the harvested area are stored. The data comprises only the logging activity, uncontrolled deforestation and degradation is not covered.

Quality control of the datasets is possible only to a limited extend, since only for some plots traces of harvesting are visible in the available VHR imagery, mostly only skit tracks. It is thus not possible to check the spatial and temporal extend of the harvested areas. For the years from 2010 onwards (and possibly earlier) the data is not complete, since according to MSD staff, not all changes have been digitized submitted. Many areas overlap in succeeding years. The number of missing values differs per year and attribute.

4.1.4 National Forest Inventory Plots

The national forest inventory of 2006 comprised a large number of plots (>1000) and reports the forest type and detailed inventory data. Spatial information was gathered with simple GPS, resulting in offsets. The plots where never revisited and thus can’t be used for change product accuracy assessment.

4.1.5 Permant Sampling Plots

The Fiji Forestry Department permanent sampling plots (PSP) were established from 2010 to 2011 on Viti Levu, Vanua Levu and Taveuni. The 50 x 50 m plots are revisited every second year, two measurement cycles were completed by end of 2013. In total there are 100 plots which follow a regular design in a 12 km (Vanua Levu) or 13.5 km (Viti Levu) grid (see Figure 4). The program is panned to continue for approximately 25 years. The recorded parameters include forest cover (no forest, mangrove, and plantation, open and closed indigenous forest). In two subplots per plot, individual trees are measured, leading to volume estimations (for details see Fiji Department of Forestry MSD 2010; Payton 2012). The data is stored in a MS Access database.

The PSP sampling plots provide a great spatial coverage, giving a probability sample of the complete main islands. However, the number of plots is too low to validate a forest cover change product. With just 100 plots, small forest-change will be underrepresented, since it typically covers only a small proportion of the monitored area. In the literature systematic designs are therefore not recommended for change product validation (Stehman 2009; Foody 2002).

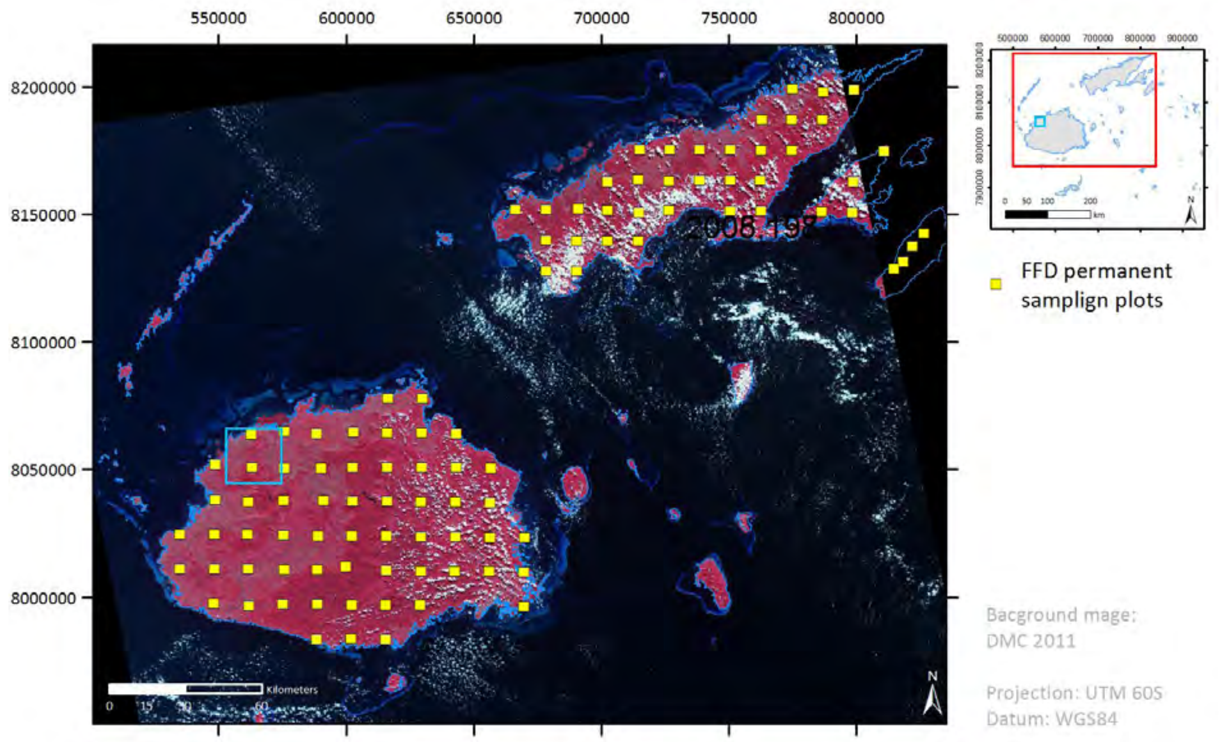


Figure 4: FFD permanent sampling plots grid

4.1.6 Digitized VHR

At present, no forest cover change datasets based on digitized VHR imagery are available.

Deriving forest change requires bi- or multi-temporal coverage. SOPAC is hoping to acquire full VHR coverage of Fiji every five years. Efforts are made towards (semi-)automated forest stratification and land cover classification. This could be used to derive forest cover change maps with fine spatial resolution. These would be very valuable for assessing spatially coarser RS-derived change products, since they provide complete surveys of the covered area. However, limiting factors are the coarse temporal resolution and data gaps due to cloud cover.

Change data derived from VHR data, typically gives a complete survey covers the spatial overlap of the used images. However, while selected for low cloud cover, there are some remaining clouds in the imagery. This leads to data gaps which sum up for multi temporal change analysis.

4.2 Database design

4.2.1 User Requirements

The database should hold spatio-temporal information on different thematic types of forest cover change for Fiji. The user should be able to i) add new forest cover change data from different sources and in different formats, ii) retrieve maps showing forest cover change for a specific temporal and spatial extend. Further, the user should be able to retrieve additional information, e.g. metadata concerning the data source and quality if available. The following list gives an overview of the requirements following a categorization adapted from Yeung and Hall (2007, chap. 3).

- Necessary Data:
    - Change data including theme (change type or no change) and spatio-temporal extent.
    - Meta data giving Information on spatial support (sampling design, original pixel size) and the data source
- Application Interface:
    - Possibility to integrate data from different sources and in different formats
    - Query for change data of a given spatio-temporal extent from GIS and analysis software.
- Performance: Performance is of minor importance, given the relatively small amount of data.

4.2.2 Conceptual Database Model

a) Definition of Database Themes: Forest Cover Change Types

In the available reference data, only a subset of the possible FCC types is covered. Additionally to the change events, there a no-change areas, in both forest and non-forest. The working definitions for the database are based on the available data; they are listed in Table XX. They are data-based and have not been assessed in the field, only by visual interpretation of VHR imagery. These definitions partially correspond to the definitions currently used by the FFD and give some information on the drivers. If the FCC definitions are further specified in the future, the change themes have to be altered accordingly.

Table 4: FCC theme working definitions, with corresponding FFD definitions and datasets that contain them.

<table>
<tr><th>FCC theme</th><th>Working definition</th><th>Corresponding FFD definition (based on FAO 2004)</th><th>Datasets</th></tr>
<tr><td>Full harvest of plantation</td><td>Areas marked as logged in FPL data</td><td rowspan="3">Potentially degradation (if logging is not sustainable)</td><td>FPL harvesting data</td></tr>
<tr><td>Selective harvest of plantation</td><td>Areas marked as logged in FHCL data</td><td>FHCL harvesting data</td></tr>
<tr><td>Selective harvest of indigenous forest</td><td>Areas marked ads harvested in MSD harvesting data, PSP plots with change in forest cover class inside harvest plots</td><td>FFD harvesting data, PSP</td></tr>
<tr><td>Replanting of plantation</td><td>Areas marked as replanted in FHCL and FPL data</td><td>Forest improvement</td><td>FPL and FHCL replanting data</td></tr>
<tr><td>Uncontrolled Deforestation of indigenous forest</td><td>Change of land cover from any forest class to non-forest in PSP data outside harvesting areas</td><td>Deforestation</td><td>PSP</td></tr>
<tr><td>Uncontrolled degradation of indigenous forest</td><td>Change of land cover from closed to open forest in PSP data outside harvesting areas</td><td>Degradation</td><td>PSP</td></tr>
<tr><td>Landslides affected plantation areas</td><td>Areas marked as landslide affected in FPL data</td><td>Degradation, potentially deforestation</td><td>FPL Landslide data, PSP</td></tr>
<tr><td>(No-change)</td><td>Without activities in plantation data, PSP with unchanged forest cover class</td><td>Forest or Non-Forest land</td><td>All datasets</td></tr>
</table>

b) Representation of time and change

A traditional way of storing temporal change is a simple snapshot approach, where for each time-point a complete map is stored. This approach is suitable for change datasets that consist of a set of repeatedly revisited sample plots, e.g. the PSP plots. For data sets that contain only the change locations (such as the plantation harvest data) this doesn't work, since the complete map is not known. Further, to represent yearly or sub-yearly changes many snapshots would be necessary, or the temporal resolution would have to be reduced. Snapshot based approaches also lead to redundant data for no change areas (Longley et al. 2005, chap. 8.3.1; Peuquet and Duan 1995).

Therefore representation of change in the database follows an entity-based approach. (Longley et al. 2005, chap. 8.3.2). The entities are the changed features. They are represented by spatial information (a multipart polygon) and a change event, which stores thematic information, including time and type of change. A change location can be linked to multiple change events.

No-change areas are represented in two ways: for datasets that comprise a complete survey of an area, no-change is stored in an indirect way through baseline features. These features serve as baseline reference for a given spatio-temporal extent. The no-change area for a baseline feature is the part of the baseline extend not covered by any change feature with change time within the baseline time interval (see Figure 5). Thus the forest cover status of every spatio-temporal location within the baseline extend is known theoretically. This approach allows retrieval of time-based representation of change resembling the event-based spatiotemporal data model introduced by Peuquet and Duan (1995).

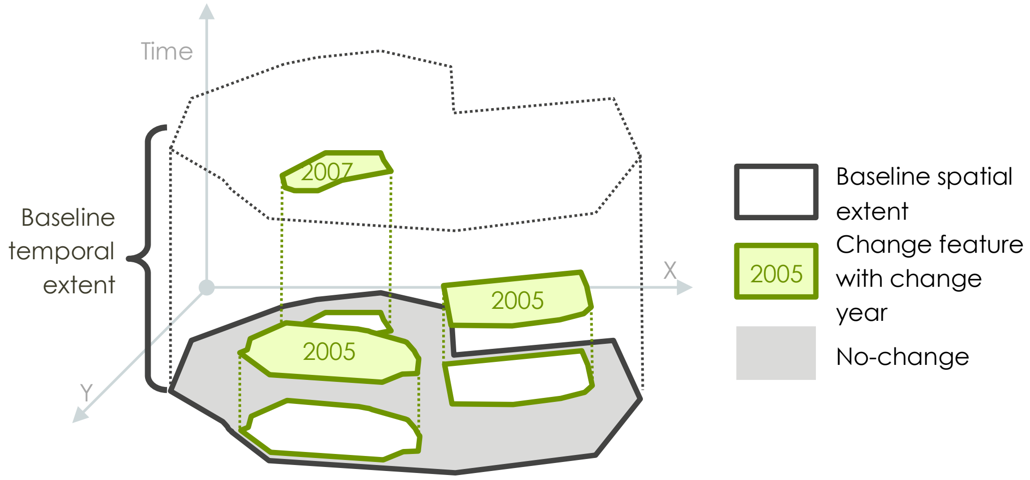


Figure 5: Concept of baseline and change features.

The second way of storing no-change areas applies to dataset with point sampling, e.g. survey plots. They typically store land-cover rather than change information. The information therefore has to be converted to change / no-change information (see section *3.3*). For simplicity, no-change is stored as special change feature. Baseline features for point-sampled dataset represent the total spatio-temporal extent that the sampling covers and for which validation thus can be conducted.

The baseline feature ID or name is used as key to match baseline features to relevant change entities. For each baseline feature, the corresponding change areas can be retrieved through an attribute query[3]. Based on this approach, a conceptual E/R model was developed (see Figure 6).

Time is stored in two formats: firstly as year of change, with additional sub-yearly information were available (quarter of the year or month). Secondly,

Time is stored as a time period with a start and end date. Additionally, year of change and sub-yearly information are stored if available, since this format is more handy for validation.

The database concept does not rule out overlapping change features. When retrieving change maps, rules must be defined to specify how such overlaps are represented in the map.

[3] For complete survey data, the relevant change features can also be retrieved by combining a spatial with a temporal (attribute) query. However, using an attribute key should allow faster processing,

The event-based approach of storing change does not explicitly store the land-cover at each location and time point. This is not required for the intended use of the database. Where information on the baseline land-cover is available, this information can be derived by defining a resulting land use for each change theme.

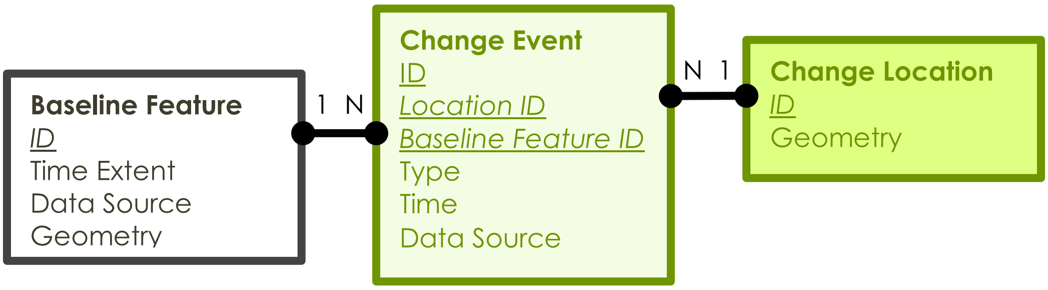


Figure 6: E/R model of conceptual FCC database design

4.3 Data Harmonization and Integration

4.3.1 GIS Forest change features

The harmonization of GIS change data requires the following steps:

- Necessary attributes are renamed and converted to comply with the change-DB format.
- Entries with missing or invalid required attributes are dropped.
- Superfluous attributes are dropped.
- The data is transformed into the database projection, if it is different.
- The change table is then updated with the harmonized dataset.
- Where possible, quality information is derived from visual comparison with VHR imagery. This involves marking change features where no change occurred, where the change status is unclear and heterogeneous plots. Spatial and temporal offsets can also be stored.

For every change feature a corresponding baseline feature has to be added to the database. The essential quality control for baseline features is to check that all forest cover change within the baseline spatial-temporal extend is not covered by a change feature. Where not, this change areas (or timespans) have to be excluded from the baseline feature or filled by other reference data (e.g. from derived from VHR imagery).

4.3.2 VHR imagery derived change data

In order to derive forest cover change data from VHR imagery, some principles should be respected: The VHR imagery has to be preprocessed carefully; this includes geometric correction and cloud masking. Change detection should not be applied on classified data (post-classification change), else the classification error of two classifications can sum up often amount of change is underestimated and its magnitude overestimated (Foody 2002). The baseline area is typically the overlap of the used images; with a temporal extend of the period between the acquisition dates.

4.3.3 PSP and Other Survey Plots

As described in *4.2.2b) Representation of time and change,* survey plots such as the PSP store land cover rather than change information. The following two steps convert the data to the change-event based database format:

1. Create change features with type “no-change” for all plots, setting the start- and end-date period to the first and last visit with the same forest cover.

2. Create change features for each plot with a change in forest cover. The period between the survey visits between which the change occurred is the change period. The change type can be derived from a look up table linking change types to the forest cover before and after the change (e.g. deforestation = change from open or closed forest to no-forest)

Note that the spatial information for the change features has to be stored only once, since multiple change events can be linked to one location.

The baseline area is the area covered by the survey. For the PSP this is the whole area of the main islands. The temporal extent is the period between the first and last survey visit.

4.4 Validation Concept

4.4.1 Concept of Accuracy Assessment

The proposed validation concept takes into account three different levels of temporal information on the detected change:

1. Bi-temporal change: The RS-based change product depicts the change between two time points.
2. Yearly change. Here we can validate a) change occurrence b) change time. Optimally, the year of change is available from both RS-based change product and the reference data. If the reference data provides only bi-temporal information (e.g. digitized VHR imagery), I might not be possible to validate the time of change.
3. Sub-yearly change: If the RS-product and reference data give more detailed information of the time of change, the change time detection can be validated more accurately.

In the bi-temporal case, the classes are simply no-change and change or different type of changes, while in the yearly scenario, there is a no-change class, and one class for each year per change type. The third scenario (sub-yearly periods) can be handled similar if the reference datasets comes with fixed periods (e.g. month). There are simply more change classes. If this produces too many classes or the reference data change periods are not fixed the sub-yearly information is difficult to use. Solution can be aggregation and reporting of the lag of change detection.

4.4.2 Accuracy Measures

The accuracy is assessed based on the confusion matrix, and a number of commonly used accuracy measures (Foody 2002). The error matrices are not normalized or standardized, since can lead to biased estimators (see e.g. Stehman and Czaplewski 1998).

The proposed accuracy measures include the overall accuracy (OA) or correct proportion, and producer's and user's accuracy per class. For point sampling, a confidence interval for the overall accuracy is calculated to capture the uncertainty introduced by sampling. Cohen's kappa coefficient is not calculated, since it is often not considered adequate for expressing land cover change accuracy (Olofsson et al. 2013; Pontius and Millones 2011). An additional accuracy indicator for yearly and sub-yearly change detection and reference data (cases 2 & 3) is the mean time lag of change detection. This is the time difference between the reference change time and the detected time. See Table 5 for interpretation and calculation of these measures.

Table 5: Proposed accuracy measures with interpretation and calculation.

Measure	Interpretation	Calculation
Confusion Matrix	Sums up correctly classified and confused samples per class.	Matrix of counts n_{ij} of samples mapped as class i with reference class j. Note that for stratified samples area corrected consistent estimators have to be used[4].
Overall Accuracy (OA)	Ratio of correct detected samples	$OA = \sum_{i=1}^{m} \frac{n_{ii}}{n}$ with m the no. of classes
Confidence interval (CI) for Overall Accuracy	Sampling uncertainty for OA. Not applicable for complete survey sampling.	Calculation depends on sampling design. For a sufficiently large simple random sample with sample size n, the CI at confidence level α can be approximated as $CI_{OA} = OA \pm t_{1-\alpha}\sqrt{\frac{OA(1-OA)}{n-1}}$ Card (1982) provides variance estimators for stratified random sampling.
Producer's accuracy (per class)	Share of the of correctly classified samples in a reference class	$\frac{n_{ii}}{n_{*i}}$; with n_{*i} being the reference class total sample count
User's accuracy (per class)	Share of the of correctly classified samples in a mapped class	$\frac{n_{ii}}{n_{i*}}$; with n_{i*} being the mapped class total sample count
Mean time lag of change detection	How average time between RS-based detection of change and reference data time of change. Not for bi-temporal change detection.	Mean of difference between detected and reference data time of change of all sample with change in both reference and TRS-based product.

For complete sampling (see above) the accuracy measures can be calculated directly from cell counts (as provide by R package caret (Olofsson et al. 2013; formulas from R package caret are used, see Kuhn 2008). For stratified random sampling the confusion matrix and accuracy measures and area estimators are calculated based on the unbiased estimator of the proportion of area (see Olofsson et al. 2013; Stehman and Czaplewski 1998).

4.5 Demo Implementation

4.5.1 Selection of Study Site and Data:

For the demo implementation, Lololo Pine plantation is chosen. For this area, good reference data as well as RS-products of the same time period are available. Forest cover change is clearly visible in both VHR and Landsat imagery. The Lololo lease area as stored in the land use data provided by SOPAC serves as baseline area. The available VHR imagery was used to visually confirm that area not covered by change polygons shows no major forest cover change.

It should be noted that forest cover change within the plantation is of minor interest from a deforestation or degradation point of view (see paragraph 4.2.2a).

[4] Area corrected consistent estimator of confusion matrix cells for stratification by class: $\hat{p}_{ij} = \frac{n_{ij}}{n_{i*}} \frac{N_{i*}}{N}$ with n_{ij}= number of samples in mapped as class i with reference class j, n_{i*}= sampled cells with map class i, and $\frac{N_{i*}}{N}$ the proportion of all cells mapped as class i (Olofsson et al. 2013).

4.5.2 Demo Database

a) Software and Hardware Choice

Main focus is to make use of existing spatial data infrastructure (SDI) and software that is available to the local partners. For processing the data ESRI ArcMap and the open source language and environment for Statistical computing R (R Development Core Team 2013) is used. ArcMap is used by SOPAC and the Forestry department MSD, and R is freely available at **http://www.R-project.org/**.

The database tables are stored as layers on a GeoNode platform hosted by SOPAC at **http://geonode.sopac.org/**. GeoNode acts as a geoserver and web based geo information system. It allows the up- and download of spatial data in multiple formats. Layers can be combined to maps and published through web services. Advantages are the easy use, availability to all partners, handling of meta-data and central hosting. However, the user interface doesn't allow access to the underlying relational database. Queries and linking tables is only possible to a limited extend.

b) Data standards and metadata

For easy exchange between the software packages, the data is kept in ESRI shape file format locally, since no geodatabase format exists that can be handled by all used software packages. Within the packages, native formats such as the sp formats in R as used (see Bivand, Pebesma, and Gómez-Rubio 2013). The demo final database is stored on the GeoNode platform internally in PostGIS format, compatible to Open Geospatial Consortium (OGC) and International Standard Organisation (ISO) geodata standards. Data can be available through OGC compatible Web Feature Services (WFS) and downloaded in multiple formats, including ESRI shapefiles.

For metadata, the GeoNode platform uses a combination of form based entry and automatic extraction (e.g. of spatial extend and projection). The metadata can be viewed online and downloaded in a following a variety of standards, including ISOTC211 meta-data standards.
For the demo implementation metadata elements are entered when uploading data to the database server, or imported from an external metadata file. A shortcoming is the handling of meta-data for exported data, since most export formats do not support the meta-data to be carried along.

c) Database Schema

The adopted physical scheme simplifies the conceptual schema by dropping the separation of spatial and thematic information for change features. This decision was made since there are no multiple change events per location in the used data, and to keep the database simpler. Table 6 shows the change feature table schema. The baseline feature schema is illustrated in Table 7. It follows the conceptual schema. In the demo database, only one baseline feature per change feature is allowed.

Table 6: Schema of change feature table. The key attributes (FID, foreign key: Baseline) are underlined.

Attribute	Data format	Domain/Format/Unit	Content	Required
FID	Integer	1 to number of change features	Unique feature identification number, key attribute	●
OrgFID	Integer		Feature Identification number in original dataset	●
OrgID	String	Various	Other Identification attributes from in original dataset, e.g. Stand/Coupe ID for Fiji Pine data	-
Type	String	See Table 4	Forest cover change type or no-change	●
Baseline	Integer	Baseline feature FID	Foreign key to match with baseline feature	●
Year	Integer	YYYY	Year of change	-
Period	String		Sub year period	-
Start	Date	YYYY-MM-DD	Starting date of change/no-change period	●
End	Date	YYYY-MM-DD	End date of change/no-change period	●
VolumeRm	Float	m^3	Volume removed, where applicable	-
AreaDiff	Float	ha	Area harvested or replanted, where applicable	-
DataSrc	String	FFD, FHCL, FPL, or SOPAC	Data Source (agency name)	●
Location	String		island, province, mataqali where available	-
Geometry	Multipart Polygon	Projection: WGS 84 UTM 60 South	The spatial information	●

Table 7: Schema of baseline feature table. The key attribute (FID) is underlined.

Attribute	Data format	Domain/Format/Unit	Content	Required
FID	Integer	1 to umber of baseline features	Unique feature identification number, key attribute	●
OrgFID	Integer		Feature Identification number in original dataset	●
OrgID	String	Various	Other Identification attributes from in original dataset, e.g. Forest attribute for Fiji Pine	-
Sample Type	String	Complete-survey random, stratified, systematic	Sampling type	●
Start	Date	YYYY-MM-DD	Starting date of sample period	●
End	Date	YYYY-MM-DD	End date of sampling period	●
DataSrc	String	FFD, FHCL, FPL, or SOPAC	Data Source (agency name)	●
Geometry	Multipart Polygon	Projection: WGS 84 UTM 60 South	The spatial information	●

Even though it's underlying relational database system, the GeoNode platform doesn't allow joins on the user interface. Therefore combining of baseline and change features is implemented in GIS or analysis scripts software.

This setup allows simple conversion between different data formats (e.g. GeoNode layers, ArcMap Geodatabase feature-sets and shapefiles, R spatial dataframes). The data can thus be exchanged between different packages, systems and users easily.

d) Interfaces

Data can be uploaded manually using the GeonNode website, or using the open source Quantum GIS. From within R, GeoNode layers can be downloaded using WFS. Spatial and attribute queries can be passed on in the download link. Unfortunately, the only format that works reliably is ESRI shapefile.

Therefore the importing R-script, downloads a zipped shapefile, unzips the data loads its using the rgdal package.

4.5.3 Data Integration and Harmonization

The reference data is cleaned by removing entries for which the change year attribute was missing or clearly in wrong format. The temporal information is converted to the start- end-date format using an R script. Based on the visual comparison with VHR and Landsat data, a quality flag is added, to indicate plots where no apparent change occurred and with unclear change status.

Attributes are renamed and added according to the change data schema described above, missing information (such as data source) are filled in from the meta-data. The baseline area is retrieved from the SOPAC data on pine plantation lease areas, and its attributes were changed and added according to the baseline data schema.

4.5.4 Demo Validation

The demo reference database is used to validate a BFAST derived change product of the Lololo study site from 2006 to 2012 (see section 2.4 *Remote Sensing Based Change Products*). Validation performed used using a) bi-temporal and b) yearly change information. The bi-temporal validation simply uses the same change product, but classifies all detected change in a single change class, the reference data is simplified accordingly. The reference data was used i) as provided and ii) filtered by the quality flag, treating areas where no change could be visually observed in the RS imagery as no-change areas. Figure 7 shows a detail of the BFAST change map with the corresponding reference data.

Table 8 shows the resulting overall accuracies per case; Figure 8 shows a detail of the result as a map. The validation of bi-temporal change detection shows generally higher accuracy. This is expected because confusion between the change years does not add to the overall error here. The filtering of the reference data by the quality flag resulted in only little higher accuracy for the yearly case (a,ii) and neglect able difference for the bi-temporal cases (b,ii vs. b,i).

Table 8: Overall accuracy per case

Case	Overall accuracy
a,i): yearly change, unfiltered reference data	0,7828
b,i) bi-temporal change, unfiltered reference data	0,8318
a,ii) yearly change, filtered reference data	0,7938
b,ii) bi-temporal change, quality filtered reference data	0,8398

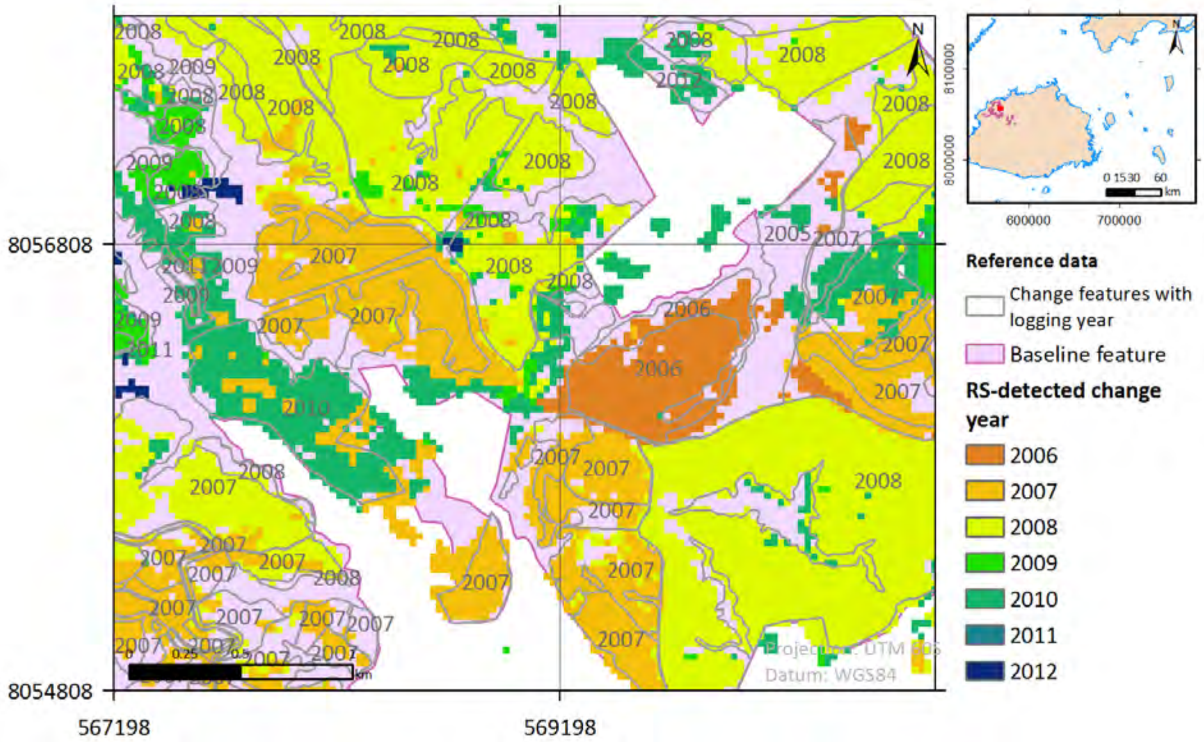


Figure 7: Demo validation RS-detected change vs. reference data (detail)

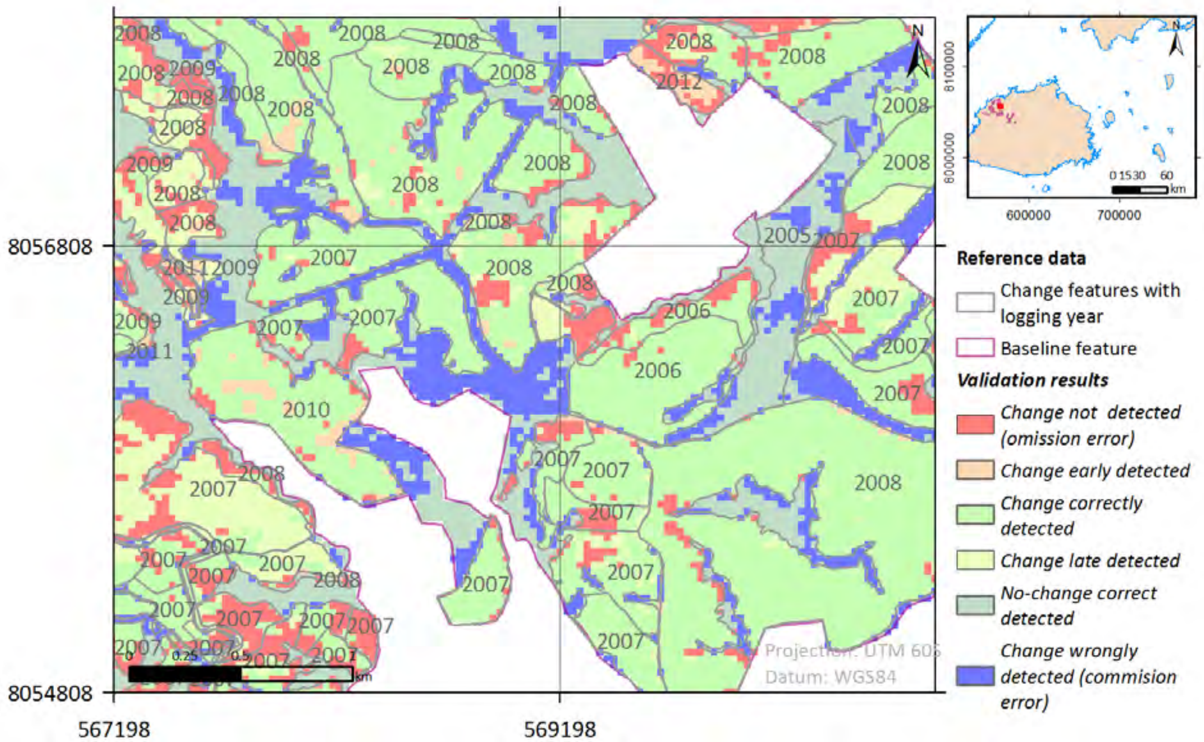


Figure 8: Results of the demo validation (detail).

Table 9 and Table 10 show the confusion matrix and per class accuracy measures respectively for case a,i). Only a relative small proportion of no-change pixels were wrongly detected, plot boundary effects can be observed. The most abundant confusion between change years is the detection of change in 2009, 2011 and 2012 as change occurring in 2010, reflected in the low user's accuracy of 2010. This can be related to the droughts occurring in the study site in 2010 (see section: *Study Site*). The detection for 2011 and 2012 is generally very bad, reflected in the low and very low producer's and user's accuracy results for those years. Apart from the misclassification due to drought, this could also be attributed to the short monitoring period with only few pictures for 2012 available. Table 11 show the confusion matrices and accuracy measures for case b,i). For the cases a,ii) and b, ii) with quality filtering, only the overall accuracy is reported, since the differences were small. Given the probable cause for the errors explained above, it is not surprising that the quality filtering of the reference data showed little result.

Table 9: Confusion matrix (pixel count) for yearly change and unfiltered reference data (case a,i)

		Reference Class								
		no-change	2006	2007	2008	2009	2010	2011	2012	Row sums
Mapped Class	no-change	120531	1126	2497	1095	1099	2978	4742	1840	135908
	2006	1396	1701	60	4	29	72	91	78	3431
	2007	1413	1058	3042	140	64	84	31	71	5903
	2008	1809	27	566	5480	423	177	25	79	8586
	2009	1722	6	8	557	3992	106	89	65	6545
	2010	7490	26	227	446	1855	3336	1015	797	15192
	2011	24	0	0	0	0	87	136	1	248
	2012	467	0	1	16	18	20	229	11	762
	Col. sums	134852	3944	6401	7738	7480	6860	6358	2942	

Table 10: : Produce's and User's accuracy for yearly change and unfiltered reference data (case a,i)

	Producer's Accuracy	User's Accuracy
no-change	0.89	0.89
2006	0.43	0.5
2007	0.48	0.52
2008	0.71	0.64
2009	0.53	0.61
2010	0.49	0.22
2011	0.02	0.55
2012	0	0.01

Table 11: Confusion matrix (pixel count) and per-class accuracy measures for case b,i): bi-temporal change with unfiltered reference data

		Reference Class		
		no-change	change	Row sums
Mapped Class	no-change	120531	15377	135908
	change	14321	26346	40667
	Col. sums	134852	41723	

	Producer's Accuracy	User's Accuracy
no-change	0.89	0.89
change	0.63	0.65

5 Discussion

5.1 Reference Data Availability, Suitability and Limitations

This research showed which sources of reference data are available for Fiji and how the data can be employed to validate RS-based forest cover change products. It also shows the limitations of the available data, especially in respect to indigenous forest.

In upcoming years, digitized VHR imagery can provide spatial very detailed and almost complete but temporally coarse forest cover and change information. The temporal resolution depends on the time intervals between the image acquisition dates. Since VHR imagery is quite costly, time intervals of several years are typical. Nonetheless, VHR data can provide spatially more complete reference datasets and is very valuable for controlling the quality of all other reference data. Forest cover change products based on the VHR imagery can serve as useful complement of temporally finer change detection products. However, VHR derived change detection products also require validation and therefor reference data.

For harvest and replanting, the ground data provided by Fiji Forestry Department and the plantation cooperation serves as good reference data. This is very valuable for method development. The data itself can serve as forest cover change data for nationwide monitoring of controlled forest cover change. Compared to the plantation data, the data for indigenous forest shows at the present stage more gaps. Controlling its quality using VHR imagery is possible only to a limited extend, since the selective logging leaves only minimal traces and regrowth is quick.

As described in section Validation Concept, for map validation it is crucial that the reference data follows a probability sampling. The harvest and replanting data is event based, leaving out areas where no controlled change occurred. For uncontrolled deforestation of indigenous forest, e.g. caused by other drivers such as expansion of agricultural areas, no ground data exists. Therefore statistical sound validation of change of the indigenous forest is not possible with the currently available data.

As mentioned above, VHR imagery can serve this purpose for some applications. However, to verify the VHR data, and if finer temporal resolution is required, additional reference data becomes necessary.

5.2 Data Harmonizing and Integration

It becomes clear, that a fully automated data integration procedure is not practicable with the many different formats of reference data. However, the described frameworks are applicable in the standard software packages.

5.3 Database Design

As the inventory has shown, the available reference data comes in different formats and depending on the sampling design, different validation concepts have to be applied. The proposed database format takes these differences into account with a simple and flexible simple and database schema. One limitations of the change event based approach are the lack of information on the land cover, since only change is stored.

With the proposed database design, retrieving spatial subsets based on attributes (such as administrative units) is possible only to a limited extend. However, this is inherent to the used datasets, since no consistent system of location attributes is used. For the purpose of change validation this is of minor importance.

5.4 Demo Implementation

The demo implementation showed how the developed concept can be put into praxis using available data infrastructure and software packages. It showed that the application does not require a real relational spatial database system. Since only few tables are necessary, the links between the tables can be established in the analysis or GIS software. Such a simple setup is easier to adopt and data conversion and exchange is requires less knowledge and resources, especially when storing the data on a central exchange platform such as the SOPAC GeoNode.

One problem is the handling of meta-data, since when exporting data from GeoNode is not automatically carried along with the data. The user has to access the meta-data separately and make sure that it is associated again with the exported data.

If the data platform could be accessed directly by all software packages, less data conversion would be necessary and meta-data could potentially automatically handled. However, for the time being no straightforward solution software for this exists that is suitable for the potential users.

The demo validation results show little response to the filtering based on the quality control. In the present case it is clear that reported error is due both to errors in the change detection and quality problems in the reference data. This indicates that more rigorous quality filtering would be useful. However, the quality filtering is limited by the requirement for probability sampling. This doesn't allow omitting problematic areas like the plot boundaries, since this could result in a bias. If probability sampling is not necessary (e.g. for method development), more rigorous filtering can be applied.

6 Recommendations

6.1 Clear Definitions

In order to monitor forest cover change, clear definitions of the forest cover change types are crucial. While the FAO definitions of deforestation and afforestation used in Fiji are precise, forest degradation and improvement still need to be clearly defined. This should take into account the purpose and practical considerations and be consistent with definitions used by the FAO and the UNFCCC. While there might be stable areas of degraded forest, degradation essentially is a process and has a temporal dimension. But unlike deforestation, this process might not be unidirectional. As suggested by Herold and Skutsch (2011) it might be more practical to treat degradation as unsustainable forest management rather than as a step towards deforestation.

6.2 Additional Permanent Sampling Plots

As described above, there is a lack of reference data for uncontrolled changes in indigenous forest cover. We therefore recommend establishing a set of dedicated additional sampling plots to complement the existing permanent sampling plots.

These plots should be selected a probability sampling design. Stratification should be done by forest class and change potential, where areas of higher change potential should be sampled more densely. Change potential could be determined by population density and infrastructure. Unfortunately, due to the even distribution of settlements, there might not be a gradient strong in those parameters. Possible exceptions are the surroundings of growing settlements. Alternatively, change potential stratification could by derived from RS-change detection or expert knowledge. However, this requires further research.

Costs can be reduced by clustering and additional stratification by accessibility. The survey could follow a much simpler procedure than in the established PSP project, since only forest cover needs to be verified. However, if additional parameters were surveyed the plots could however serve multiple purposes.

Using the plot locations of the 2006 national forest inventory is not recommended, since their sampling design is not stratified.

6.3 Further Investment in Very High Resolution Imagery

VHR satellite imagery can contribute as spatial complete, accurate and fine reference data. This makes it veryy valuable for validation of scale forest cover change especially degradation. The only for change validation is the coarse temporal resolution, due to high acquisition costs and cloud cover. It may therefore not be used for validation of yearly change monitoring, but can serve verification purposes. Further the data can serve other purposes so as e.g. forest stratification and land cover mapping.

6.4 Central and Transparent Data management

In order to make the validation process transparent and reproducible, data management should be considered carefully.

Currently important dataset are sometimes stored only locally on a single system. It is recommendable to use a central storage system with regular backups. In this way it can be ensured that every user always uses the most recent version of a dataset and the risk of data loss can be reduced.

Data should be stored or backed up in common industry formats, to decrease dependence on a specific software package.

Meta-data should be used to document the data origin and format and to track changes and updates as described in the demo implementation. This ensures that the data format will be understandable to future or external users.

If the forest cover change product validation is part of a REDD+ monitoring, reporting and verification system, these will cover only the basics and more specific data management requirements will have to be fulfilled.

6.5 Capacity Building

The GIS, RS, statistics and data management skills required for setting up and maintaining a reference database for FCC should be build up within the implementing agency. This may require hiring an expert for a transition period. In this way it can be ensured that the reference database stays operational and it can be flexible addicted to changing validation demands.

Literature

Asner, Gregory P. 2011. "Painting the World REDD: Addressing Scientific Barriers to Monitoring Emissions from Tropical Forests." *Environmental Research Letters* 6 (2): 021002. doi:10.1088/1748-9326/6/2/021002.

Bivand, Roger, Tim Keitt, and Barry Rowlingson. 2013. *Rgdal: Bindings for the Geospatial Data Abstraction Library*. http://CRAN.R-project.org/package=rgdal.

Bivand, Roger, Edzer Jan Pebesma, and Virgilio Gómez-Rubio. 2013. *Applied Spatial Data Analysis with R*. New York, NY: Springer.

Card, Don H. 1982. "Using Known Map Category Marginal Frequencies to Improve Estimates of Thematic Map Accuracy." *Photogrammetric Enginneering and Remote Sensing* 48 (3): 431–39.

Chaudhary, Felix. 2012. "Blame It on the Rain." *Fiji Times Online*, February 8, sec. Front Page News. http://www.fijitimes.com/story.aspx?id=192818.

Congalton, Russell G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91)90048-B.

De Sy, Veronique, Martin Herold, Frédéric Achard, Gregory P Asner, Alex Held, Josef Kellndorfer, and Jan Verbesselt. 2012. "Synergies of Multiple Remote Sensing Data Sources for REDD+ Monitoring." *Current Opinion in Environmental Sustainability* 4 (6): 696–706. doi:10.1016/j.cosust.2012.09.013.

Encyclopedia Britannica. 2014. "Fiji (republic, Pacific Ocean)." *Encyclopedia Britannica*. http://www.britannica.com/EBchecked/topic/206686/Fiji.

FAO. 2004. "Global Forest Resources Assessment 2005 – Terms and Definitions (Final Version)". Working Paper 83. Rome: FAO. http://www.fao.org/forestry/9687-0f7ba44a281b061b9c964d3633d8bf325.pdf.

———. 2010. "Global Forest Resources Assessments (FRA)". FAO. http://www.fao.org/forestry/fra/en/.

FCPF, 2014. "Readiness Preparation Proposal (R-PP) for Fiji", Forest Carbon Partnership Facility, https://www.forestcarbonpartnership.org/sites/fcp/files/2014/February/Fiji_R-PP_rev_2014_01_22.pdf

Fiji Department of Forestry. 2011. *Fiji REDD-plus Policy: Reducing Emissions from Deforestation and Forest Degradation in Fiji*. Suva, Fiji Islands: Secretariat of the Pacific Community.

Fiji Department of Forestry MSD. 2010. "Permanent Sample Plots for Forest Monitoring in Fiji". Pacific GIS/RS Conference December, Suva, Fiji Islands. http://www.sopac.org/sopac/gis.conf.2010/Day02/Session06/Presentation_102_Permanent%2520Sample%2520Plots%2520for%2520Forest%2520Monitoring%2520in%2520Fiji_Akosita_Lewai.ppt&prev=/search%3Fq%3Dfiji%2Bpermanent%2Bsample%2Bplots%26client%3Dfirefox-a%26hs%3DFGN%26rls%3Dorg.mozilla:de:official%26channel%3Dfflb.

Foley, Jonathan A., Ruth DeFries, Gregory P. Asner, Carol Barford, Gordon Bonan, Stephen R. Carpenter, F. Stuart Chapin, et al. 2005. "Global Consequences of Land Use." *Science* 309 (5734): 570–74. doi:10.1126/science.1111772.

Foody, Giles M. 2002. "Status of Land Cover Classification Accuracy Assessment." *Remote Sensing of Environment* 80 (1): 185–201. doi:10.1016/S0034-4257(01)00295-4.

Goetz, Scott J., Alessandro Baccini, Nadine T. Laporte, Tracy Johns, Wayne Walker, Josef Kellndorfer, Richard A. Houghton, and Mindy Sun. 2009. "Mapping and Monitoring Carbon Stocks with Satellite Observations: A Comparison of Methods." *Carbon Balance and Management* 4 (1): 2. doi:10.1186/1750-0680-4-2.

Herold, Martin, and Margaret Skutsch. 2011. "Monitoring, Reporting and Verification for National REDD + Programmes: Two Proposals." *Environmental Research Letters* 6 (1): 014002. doi:10.1088/1748-9326/6/1/014002.

Houghton, R. A. 2012. "Carbon Emissions and the Drivers of Deforestation and Forest Degradation in the Tropics." *Current Opinion in Environmental Sustainability* 4 (6): 597–603. doi:10.1016/j.cosust.2012.06.006.

IPCC. 2006. "IPCC Guidelines for National Greenhouse Gas Inventories". Hayama, Japan: IGES.

Kuhn, Max. 2008. "Building Predictive Models in R Using the Caret Package." *Journal of Statistical Software* 28 (5): 1–26.

Lal, Rancha. 2013. "Pine Industry Bounces Back with Renewed Confidence." *Fiji Sun Online*, November 23, sec. Business. http://www.fijisun.com.fj/2013/11/23/pine-industry-bounces-back-with-renewed-confidence/.

Leslie, Alfred, and Osea Tuinivanua. 2010. "Fiji Forestry Outlook Study". Working Paper No. APFSOS II/ WP/ 2009/ 20. Bankog: FAO (Regional offfice for Asia and the Pacific). www.fao.org/docrep/014/am615e/am615e00.pdf.

Longley, Paul, M F Goodchild, D J Maguire, and D W Rhind. 2005. *Geographical Information Systems: Principles, Techniques, Management, and Applications*. New York; Chichester: Wiley.

Maniatis, Danae, and Danilo Mollicone. 2010. "Options for Sampling and Stratification for National Forest Inventories to Implement REDD+ under the UNFCCC." *Carbon Balance and Management* 5 (1): 9. doi:10.1186/1750-0680-5-9.

Masek, J. G., E. F. Vermote, N. Saleous, R. Wolfe, F. G. Hall, F. Huemmrich, F. Gao, J. Kutler, and T. K. Lim. 2012. *LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code*. Version.

Olofsson, Pontus, Giles M. Foody, Stephen V. Stehman, and Curtis E. Woodcock. 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* 129 (February): 122–31. doi:10.1016/j.rse.2012.10.031.

Payton, Ian j. 2012. "Development of a National Methodology for Forest Carbon Stock Assessment in FIji". Suva, Fiji Islands: Fiji Forestry Department.

Peuquet, Donna J., and Niu Duan. 1995. "An Event-Based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data." *International Journal of Geographical Information Systems* 9 (1): 7–24. doi:10.1080/02693799508902022.

Pontius, Robert Gilmore, and Marco Millones. 2011. "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment." *International Journal of Remote Sensing* 32 (15): 4407–29. doi:10.1080/01431161.2011.552923.

R Development Core Team. 2013. *R: A Language and Environment for Statistical Computing* (version 2.15.3). R Development Core Team. http://www.R-project.org/.

Rina, Samantha. 2010. "Brace for the Worst Drought." *Fiji Times Online*, September 13, sec. Fiji News. http://www.fijitimes.com/story.aspx?id=155834.

Romijn, Erika, Martin Herold, Lammert Kooistra, Daniel Murdiyarso, and Louis Verchot. 2012. "Assessing Capacities of Non-Annex I Countries for National Forest Monitoring in the Context of REDD+." *Environmental Science & Policy* 19–20: 33–48. doi:10.1016/j.envsci.2012.01.005.

Simula, Markku. 2009. "Towards Defining Forest Degradation: Comparative Analysis of Existing Definitions". Working Paper 154. Rome: FAO. http://webcache.googleusercontent.com/search?q=cache:jhj8lQzZ7g4J:ftp://ftp.fao.org/docrep/fao/012../k6217e/k6217e00.pdf+&cd=1&hl=de&ct=clnk&client=firefox-a.

Stehman, Stephen V. 2009. "Sampling Designs for Accuracy Assessment of Land Cover." *International Journal of Remote Sensing* 30 (20): 5243–72. doi:10.1080/01431160903131000.

Stehman, Stephen V., and Raymond L. Czaplewski. 1998. "Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles." *Remote Sensing of Environment* 64 (3): 331–44. doi:10.1016/S0034-4257(98)00010-8.

UNFCCC. 2013. "Decision -/CP.19 - Modalities for National Forest Monitoring Systems." In *Decisions Adopted by COP 19/ CMP 9*. Warsaw. http://unfccc.int/files/meetings/warsaw_nov_2013/decisions/application/pdf/cop19_redd_finance.pdf.

Verbesselt, Jan, Achim Zeileis, and Martin Herold. 2012. "Near Real-Time Disturbance Detection Using Satellite Image Time Series." *Remote Sensing of Environment* 123 (August): 98–108. doi:10.1016/j.rse.2012.02.022.

West, Matthew. 2011. *Developing High Quality Data Models*. Elsevier.

Yeung, Albert K. W., and G. Brent Hall. 2007. *Spatial Database Systems: Design, Implementation and Project Management*. Springer.

Zhu, Zhe, and Curtis E. Woodcock. 2012. "Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery." *Remote Sensing of Environment* 118: 83–94.