Review of the Historical Change and Classification of Forest Areas of Ecuador

Date: December 20th, 2013

Ricardo Aravena and Gabriel Eickhoff
Forest Carbon International
Vientiane • Jakarta

Forest Carbon
Mitigating Climate Change Through Conservation & Sustainable Forestry

Table of Contents

Background 3

Methods 5

Accuracy Assessments of Historical Classifications 5

2000 and 2008 Reference Imagery Selection 6

2012 Reference Imagery Selection 6

Accuracy Assessment Sample Design 7

.... 12

Revision of Historical Maps 14

Gap Filling 14

Restratification and Revision of 2012 Sampling 14

Original 2012 Sampling Strategy 14

Restratification 15

Results and Discussion 17

Change Analysis 17

Estimate of Deforestation 17

Validation of the Sampling Methodology 17

Applying the Sampling Stratification to the Previous Time Periods 17

Comparison of results with the original classification before the gap-filling 18

Conclusions and Recommendations 19

Sampling Methodology 19

Sources of Potential Error 19

Historical Mapping 20

Other Recommendations and Suggestions 20

Further validation 20

Software suggestions and capacity building for future efforts 21

Annex I: Forest Definition 22

Annex II: Equations Used to Calculate the Sample Size 23

Annex III: Equations Used to Estimate the Rate of Deforestation 24

Annex IV: Land Cover Classification 25

Background

Over the past few years the Government of Ecuador and many of its stakeholders have begun exploring new technical and policy options for the implementation of a national approach on the emerging international climate change mechanism, Reducing Emissions from Deforestation and Degradation as well as carbon stock enhancements (REDD+). At the forefront of this work in Ecuador is the Ministry of Environment (MAE) due to its mandate in regards to the administration and management of it's protected areas and the national forest estate. More recently, the MAE has been tasked with the preparation and planning for of a number of programs related to the development of a national-level REDD+ strategy.

As part of this work, MAE has been engaged by the GIZ/KfW REDD Early Movers programme (REM) along with the Government of Norway, who are interested in supporting early action under REDD+, as early as 2013. In order to do this, it was necessary to put immediate technical focus on an assessment of historical land cover change in the country, as this forms the core basis for performance-based support going forward.

Previously, the MAE had generated land cover classifications for the years 1990, 2000 and 2008. In order to rapidly respond to some of the requirements of REM with the reporting requirements set forth by the REM program it was necessary to also generate an estimate of forested land cover for 2012. Due to time and image availability constraints, it was decided to forego a wall-to-wall classification of forest cover and instead estimate forest cover by classifying a representative sample of Ecuador's 2012 land area.

In August 2013 Forest Carbon joined to support the effort already underway by MAE, the FAO and Carbon Decisions International (CDI). Forest Carbon's initial role at that time was to provide an independent review of the sampling design used to estimate the 2012 forest cover. The results of this work found that due to high rates of cloud cover and other coverage gaps, the initial sampling model would benefit marginally from revisiting the number of deforestation-based strata used to calculate the number of 2012 tiles sampled, but more importantly found that the method used for assessing the accuracy of the land cover classification for the previous years (1990, 2000, 2008) was not an approach typically used for this type of assessment and produced accuracy results below internationally recognized thresholds for jurisdictional and nested REDD+ programmes. A full explanation and description of this work is available in Eickhoff et al. 2013[1].

As a result of this report, Forest Carbon continued on with the MAE/FAO/CDI team to support MAE in undertaking an accuracy assessment based off of high resolution imagery and to work together with MAE and FAO to review and improve any geographic areas or land cover types in the 2000 and 2008 maps needing further revision in order to bring accuracy above a 75% threshold, with an aim of reaching 80%.

The work had four main objectives:

1) To review the accuracy of existing land cover classifications using high resolution-based accuracy assessment approach for each historical year (2000, 2008 and individual sampling tiles of 2012)
2) To provide backstopping and technical support to MAE with respect to any additional revisions required to the 2000, 2008 and 2012 classifications in order to bring each to a minimum accuracy of 75%.

[1] Eickhoff, G., Ferrand, J., and I. Cummins 2013. Review of Sampling Methodology for Ecuador MRV and Cost-Benefit Comparison with annual Wall-to-Wall Mapping. Internal Report to the Ecuador Ministry of Environment and GIZ REDD Early Movers Programme.

3) Support MAE in the restratification of deforestation, taking into consideration the percentage of non-cloud covered area in each sample tile and confirm that the number of resulting 2012 sampled tiles meets targeted Standard Error and Confidence Intervals for that year.

4) To provide a calculation of gross forest cover change over the 2000-2012 period.

This report outlines the work under taken, results and a discussion on the interpretation of the results given the datasets used and some possible implications for performance-based payments to Ecuador.

Methods

Accuracy Assessments of Historical Classifications

The first steps in the review and revision process was to better understand the level of accuracy associated with the existing MAE forest/non-forest classifications generated from Landsat Thematic Mapper (TM) data over the years of 2000, 2008 and 2012. The interpretation of satellite imagery and its classification into various land cover classes is a process that produces a land cover map that has a certain level of accuracy when compared to the "real-world". The use and the meaningfulness of such maps can only be known after a standard accuracy assessment has been conducted.

Previous accuracy assessment conducted on these maps used an alternative method for assessing accuracy whereby an independent remote sensing operator conducted their own wall-to-wall classification for 2000 and 2008. However, instead of using known or high-resolution imagery informed sampling points, the operators own classification was cross tabulated point for point with the MAE classification. On review it was found that the achieved accuracies were in the low to mid 70s. [2]

Table 1: Summary of previous accuracy assessments conducted by MAE.

Option	Type	Score	Qualifies[3]
Accuracy of Forest/Non-Forest			
1990	Percent Pixel Agreement	64%	No
2000	Percent Pixel Agreement	71%	No
2008	Percent Pixel Agreement	74%	No
2012	None	N/A	No
Accuracy of Change Detection			
1990	None	N/A	No
2000	None	N/A	No
2008	None	N/A	No
2012	None	N/A	No

There are two primary components of error in thematic maps such as land cover maps; positional error and thematic error. The images that were used by the MAE had already been orthorectified, so positional error was not a problem, however it was the thematic error that needed to be checked. The list of thematic land cover definitions used for this study follows the MAE classification system and is available in Annex I: Forest Definition.

Accuracy assessments are ideally conducted in association with field-level ground-truthing, but such activities are often subject to time and budget constraints. It is sufficient to use high-resolution imagery alone in place of ground truthing when conducting an accuracy assessment. Figure 1 below explains the workflow involved in the verification of accuracy when using high-resolution imagery.

In order to achieve this, satellite imagery of a higher resolution than that which was used for the original classification (Landsat 5/7) for the equivalent years was required. The selection and the location of the high-resolution imagery must first ensure that enough area will be covered to generate enough control points. Secondly it must cover as much as possible the various landscapes so that the whole classification is tested.

[2] No additional data or methods on this process were available from MAE to confirm or further examine this method.
[3] As evaluated against international jurisdictional REDD+ guidance (i.e. VCS-JNR) which require a forest/non-forest accuracy of at least 75%.

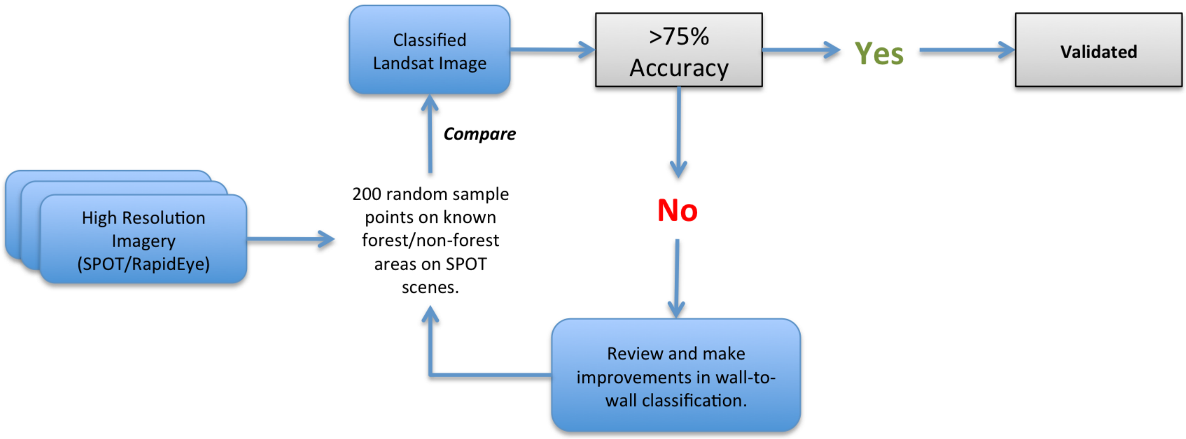

Figure 1: Accuracy assessment workflow

2000 and 2008 Reference Imagery Selection

MAE possessed nearly wall-to-wall coverage of RapidEye imagery for 2012, however high-resolution imagery for 2000 and 2008 was not available and needed to be additionally procured. 10m resolution SPOT-4 and SPOT-5 images were specifically purchased and used for the assessment of 2000 and 2008 imagery (see Table 2). Since 2000 and 2008 wall-to-wall coverage would have been prohibitively expensive and due to a scarcity of cloud-free images in the SPOT Catalog archives, 3 representative scenes were procured for this purpose for both 2000 and 2008, for a total of 6 scenes, each covering an area of 360km^2, for a total coverage of approximately 10,800km^2. In reality the coverage was somewhat less due to cloud coverage and minor overlapping between images in the 2000 imagery. While a larger geographic distribution of images for 2000 would have been ideal, almost every other image around that date contained heavy cloud cover or was located on top of high mountainous regions where little or no forest cover would likely be found. This imagery was then used as a basis for point-on-point comparisons to the historical imagery to confirm, with higher accuracy, the correct interpretation and distribution of forest and non-forest land cover. Figure 2 shows the geographic locations of the images that were finally selected.

Table 2: List of procured imagery used for accuracy assessment of 2000 and 2008 classifications.

Satellite	Product	K/J	Date	% Cloud	Usage
SPOT4	10 m color	640/350	4/11/00	4%	Accuracy Assessment
SPOT4	10 m color	640/350	4/11/00	4%	Accuracy Assessment
SPOT4	10 m color	640/352	10/7/99	2%	Accuracy Assessment
SPOT5	10 m color	640/353	24/7/07	31%	Accuracy Assessment
SPOT5	10 m color	639/358	20/11/07	48%	Accuracy Assessment
SPOT5	10 m color	643/350	12/9/08	12%	Accuracy Assessment

2012 Reference Imagery Selection

For 2012, the process was slightly different. 5m-resolution RapidEye imagery previously procured and made available by the MAE was used for the accuracy assessment. However, since the 2012 classification was conducted as a stratified sample, only certain areas of the country were available for assessing accuracy. Further, the 2012 RapidEye imagery coverage of Ecuador included only some areas of the country. Figure 3 shows the sample cells on the left and the available RapidEye imagery within a threshold or +/- 1 year. These were intersected in order to obtain the area that could be assessed for classification accuracy.

Accuracy Assessment Sample Design

A number of sampling techniques can be used, including systematic, random and stratified random. Only truly random sample designs can guarantee an unbiased sample. To ensure consistency and objectivity across the 3 coverage years, simple random sampling of each year was undertaken. Some differences in the calculated number of sample points emerged as different years required different areas to be sampled with different total areas, due to a lack of image availability for the same areas. The locations and extent of sample points for 2000, 2008 and 2012 are illustrated in Figure 4, Figure 5 and Figure 6 respectively.

For the sampling, it was necessary to generate random points across the extent of each high-resolution image for each year. Each point is visually identified as either forest or non-forest based on what can visually be seen in the high-resolution imagery, in the absence of any classification. It is assumed that the points visually identified by high-resolution imagery are an “accurate” representation of that specific pixel. Once this is completed, the visually classified points are then overlaid atop the lower resolution Landsat-based land cover classification and cross-checked for either a correct or incorrect thematic classification.

For each year, approximately 100 random points were applied to each scene. Since the number of points is a function of “visible area” the exact number of points per scene varied somewhat due to cloud cover and overlapping images. At least 4.5 km of distance was set as a sampling buffer between each point.

Once all accuracy sample points have been cross checked, “hits” and “misses” are tallied into a matrix for each year. Complete results are available in Table 3, Table 4 and Table 5. A confusion matrix numerically represents the difference between the actual and predicted classifications of a model and ultimately provides the overall accuracy of the thematic classification.

Results from the sampling indicate clearly that the accuracy of the areas sampled from the original 2000, 2008 2012 wall-to-wall classifications returned overall accuracies of higher than 90% indicating that the supervised classifications undertaken in those years were of a high enough quality to move forward with additional analyses.

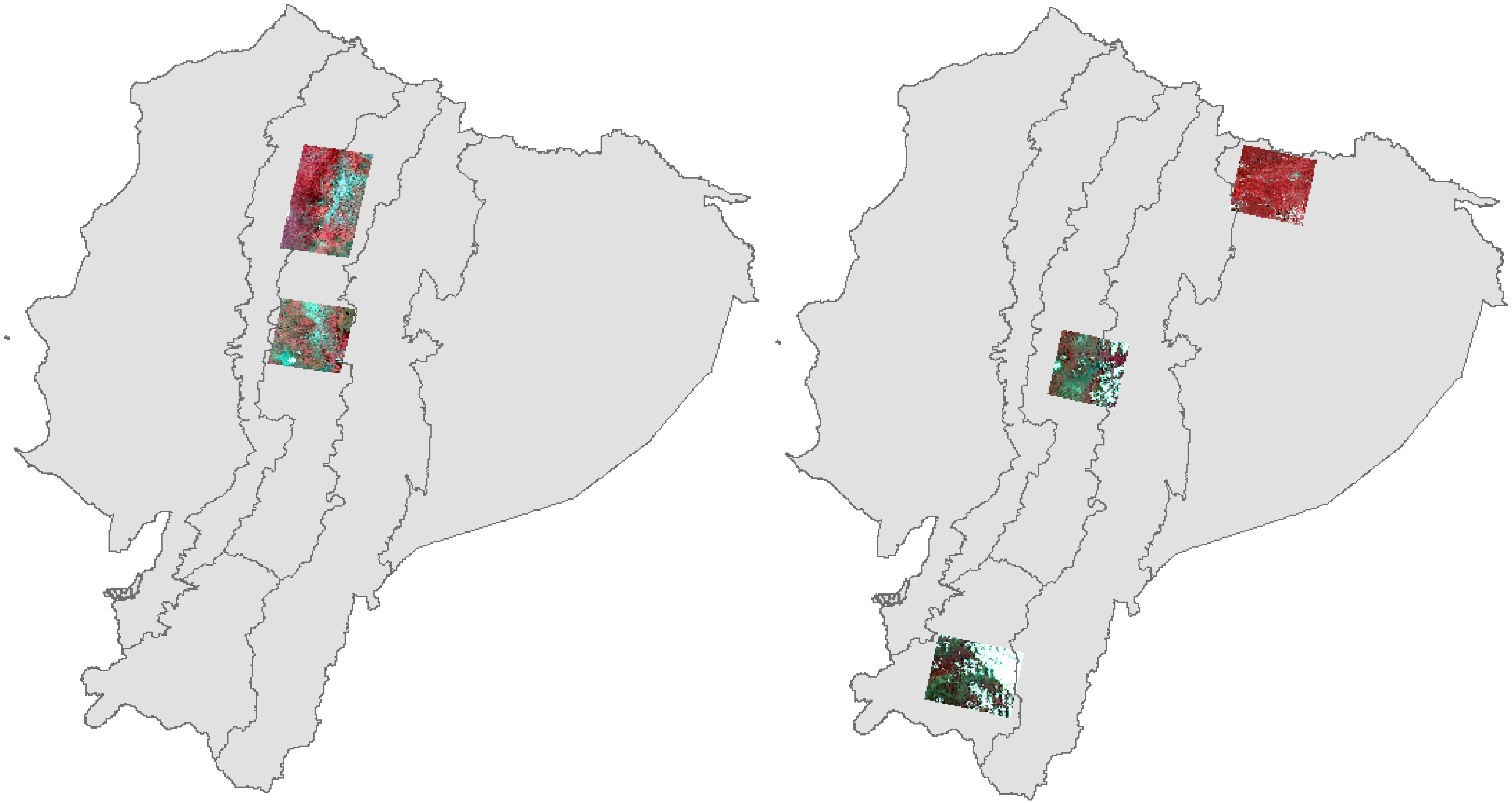

Figure 2: Accuracy assessment images for the years 2000 (left) and 2008 (right)

Forest Carbon
Mitigating Climate Change Through
Conservation & Sustainable Forestry

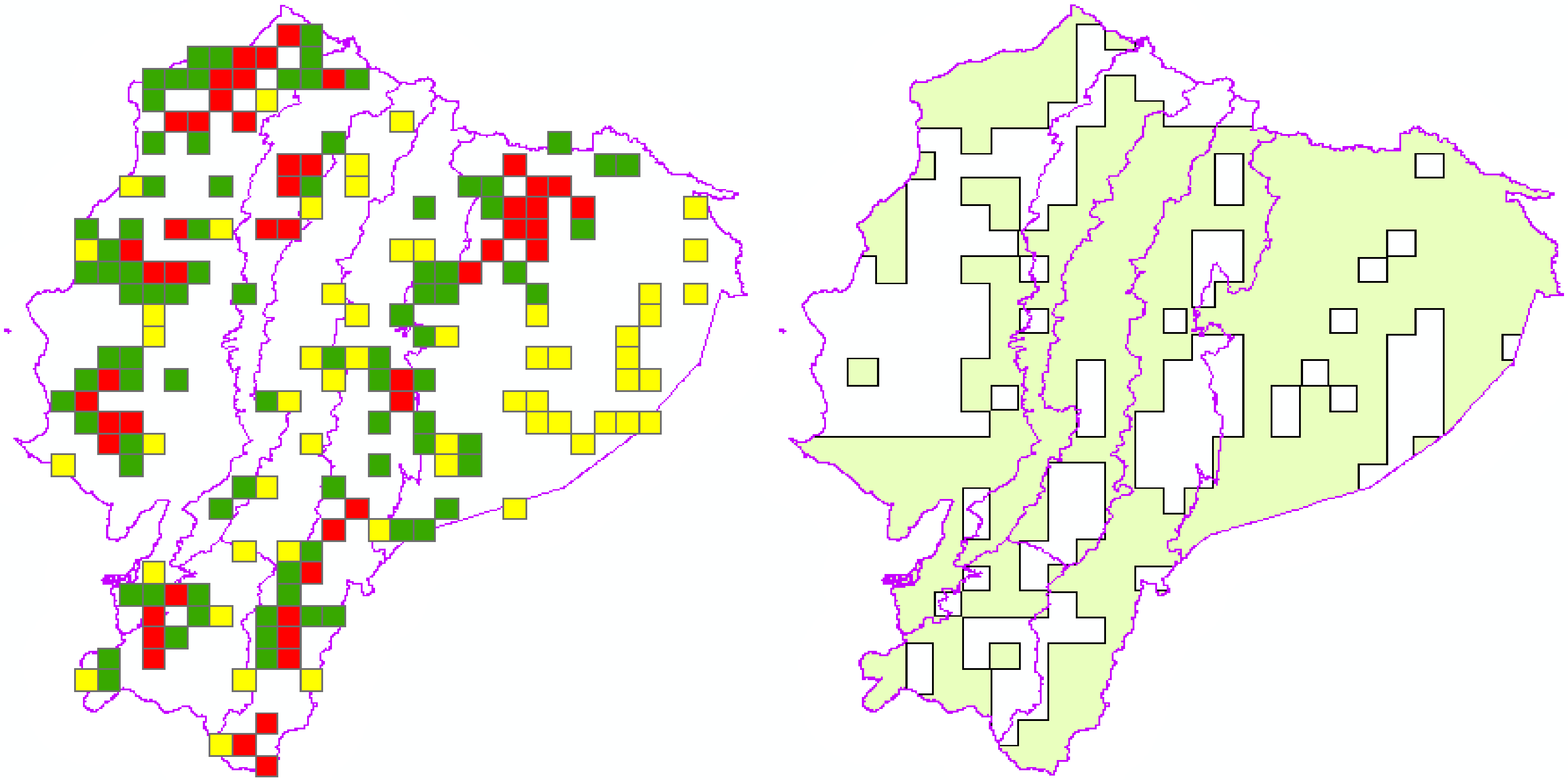

Figure 3: Intersection between stratified sample cells inside the country boundaries (left). Red, yellow and green indicate high, medium and low strata levels of historical deforestation. To the right is a representation of the available Rapid Eye coverage in green.

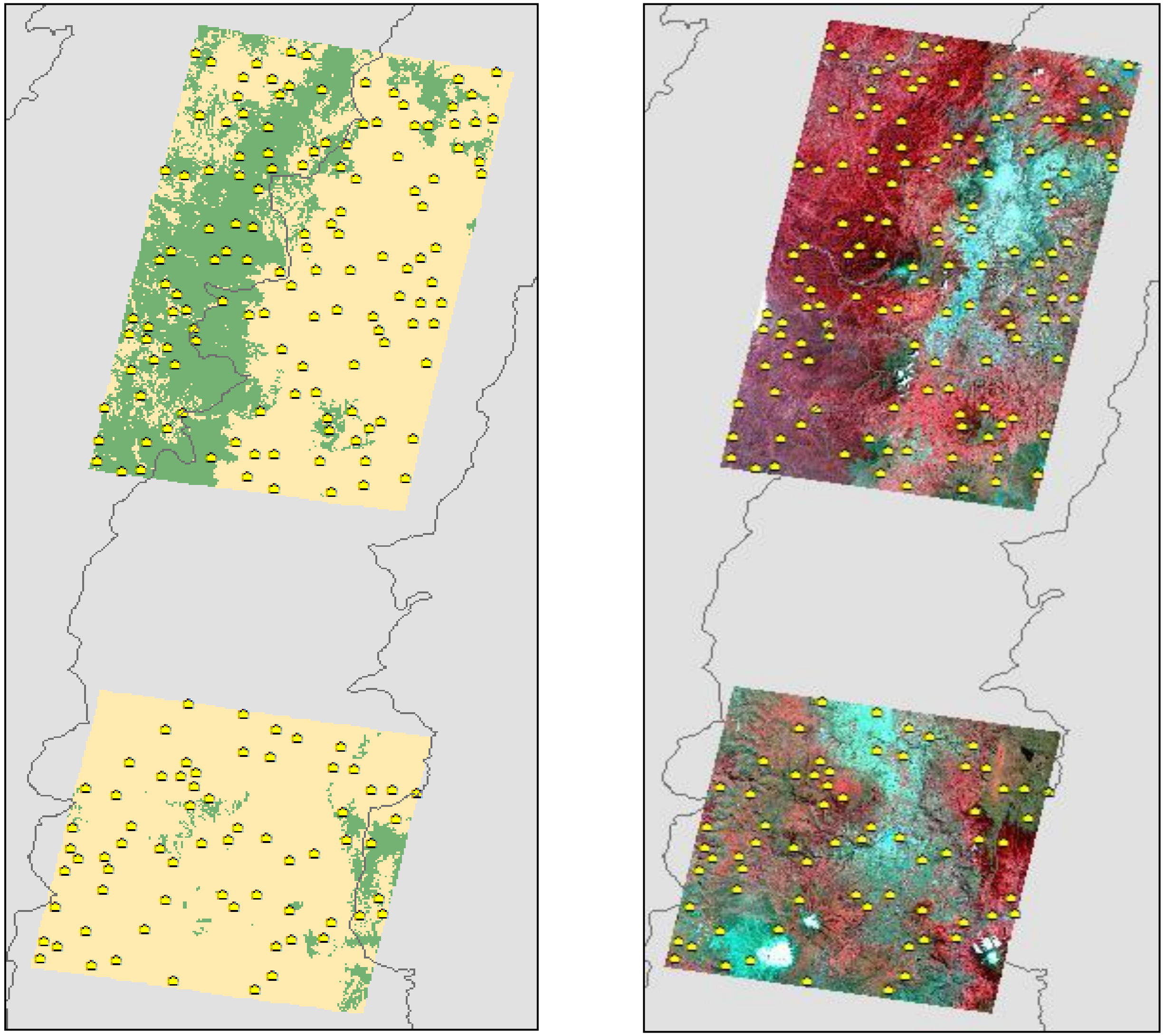

Figure 4: 2000 Landsat, 30m resolution classifications (left) assessed by 2000 SPOT-4, 10m resolution images (right) with 200 sample points

Figure 5: 2008 Landsat, 30m resolution classifications (left) assessed by 2008 SPOT-5, 10m resolution images (right) with 300 sample points

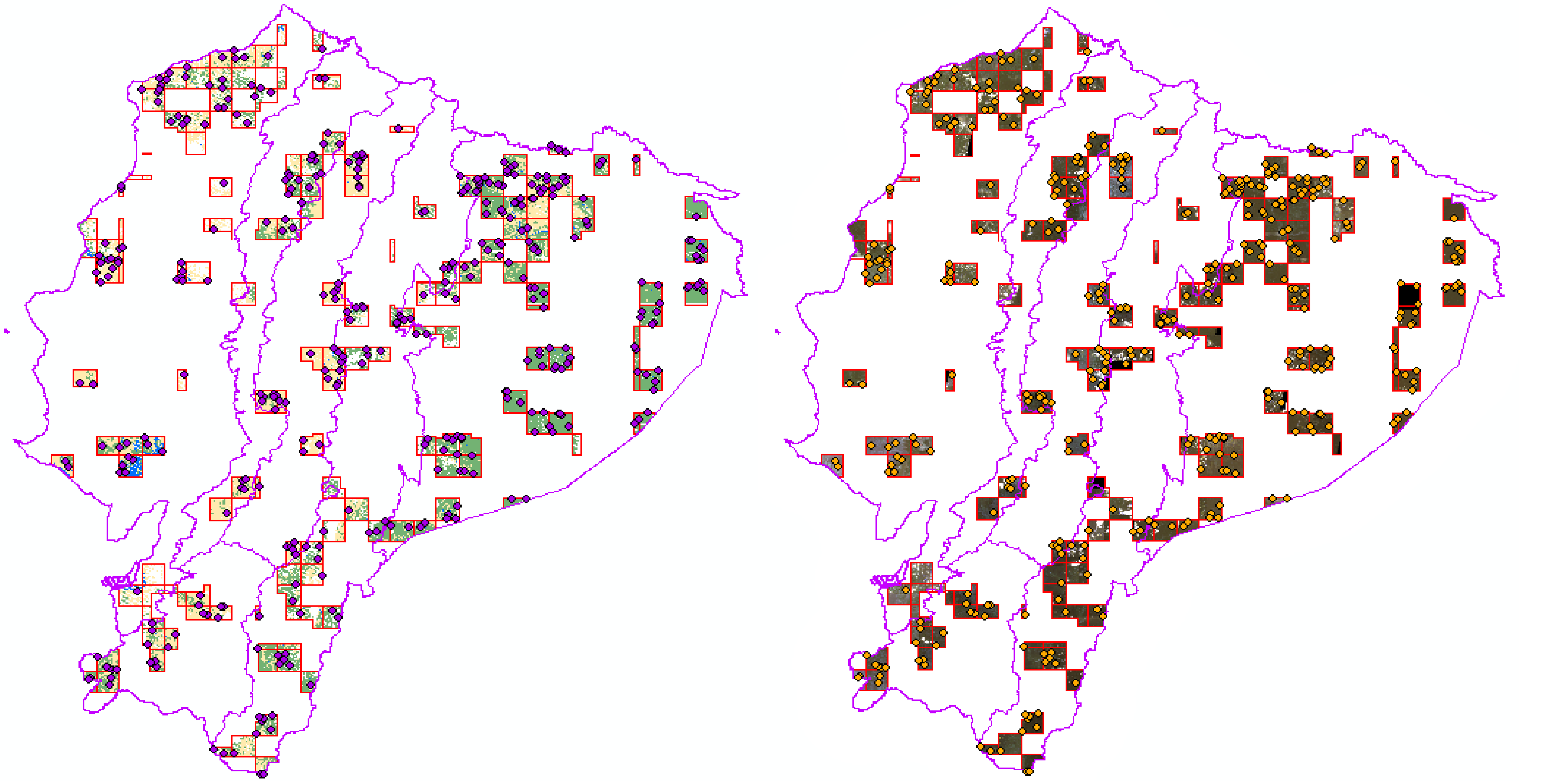

Figure 6: 2012 Landsat, 30m-resolution classifications (left) assessed by 2012 RapidEye, 5m-resolution imagery (right) with 384 sample points.

A confusion matrix was used to show the accuracy of the classification results for each year by comparing the classification results with the corresponding higher resolution imagery. This involves producing overall accuracy from producer and user accuracies, which in turn produce a kappa coefficient. Producer[4] and user[5] accuracy are determined before producing an overall accuracy. The overall accuracy is determined by dividing the total number of correct pixels (diagonal) by the total number of pixels in the error matrix.

The following are the results for each year:

Table 3: Accuracy assessment for the year 2000

Confusion Matrix: 2000		SPOT		*Classified (Landsat) pixels*	*User's Accuracy (%)*
		Forest	Non-forest		
Landsat classification:	Forest	**46**	3	49	93.9
	Non-forest	6	**145**	151	96
No. ground truth (SPOT) pixels		52	148	**200**	
Producer's Accuracy (%)		88.5	98		**95.5**

Overall accuracy = 95.5%
Average Accuracy = 93.25 %
Kappa coefficient = 88.1%

Table 4: Accuracy assessment for the year 2008

Confusion Matrix: 2008		SPOT		*Classified (Landsat) pixels*	*User's Accuracy (%)*
		Forest	Non-forest		
Landsat classification:	Forest	**100**	8	108	92.59
	Non-forest	10	**182**	192	94.79
No. ground truth (SPOT) pixels		110	190	**300**	
Producer's Accuracy (%)		90.9	95.8		**94**

Overall accuracy = 94%
Average Accuracy = 93.35 %
K = 87%

Table 5: Accuracy assessment for the year 2012

Confusion Matrix: 2012		RapidEye		*Classified (Landsat) pixels*	*User's Accuracy (%)*
		Forest	Non-forest		
Landsat classification:	Forest	**217**	7	224	96.9
	Non-forest	16	**160**	176	90.9
Ground truth (RapidEye) Pixels		233	167	**400**	
Producer's Accuracy (%)		93.1	95.8		**94.3**

Overall accuracy = 94.3%
Average Accuracy = 94.5 %
K = 91.44%

The Kappa Coefficient is the proportion of agreement after chance agreement has been removed. If kappa=1, there is perfect agreement. If kappa=0, the agreement is the same as would be expected by chance. The stronger the agreement, the higher the value of kappa, whereas negative values occur when agreement is weaker than expected by chance, but this rarely happens. Depending on the application,

[4] Producer accuracy is a reference-based accuracy that is computed by looking at the predictions produced for a class and determining the percentage of correct predictions. It represents how well a certain area can be classified (omission error).
[5] User accuracy is a map-based accuracy that is computed by looking at the reference data for a class and determining the percentage of correct predictions for these samples. It provides a measure of reliability, or the probability that a pixel class on the map represents the category on the ground (commission error)

kappa less than 70% may indicate that the measurement system needs to be improved. Kappa values greater than 90% are generally considered very good.

Revision of Historical Maps

Any years for which accuracy assessments fell below a threshold of 75%[6] were to be reviewed, revised and/or manually patched and re-classified in order to obtain a classification of greater than 75%. Given that all of the years were beyond the 75% threshold, it was not necessary to do this.

Gap Filling

While the results from the MAE classification were technical sound, a separate on-going process by MAE to fill in and complete the clouded areas of the historical maps was underway as a parallel process. Ecuador is characterized by year-round high levels of cloud cover pronounced on the border between the oceanic coast and the east of the Andean sierra. Areas without information due to the presence of clouds, cloud shadows and sensor "banding" in the satellite images are a problem because they result in a considerable loss in the amount of information available.

It was agreed by MAE to address concerns of under sampling due to cloud cover and scanner line correction error was addressed by gap filling using ancillary images. To reduce the negative effects of this problem, an attempt was made to fill the gaps with information from other images that undergo a process of radiometric adjustment by linear regression of the digital values. The selection of images that are used as gap fillers is based on the number of visible pixels available. This process was done with the OpenForis Toolkit. The classification itself was performed using a multi-temporal decision tree methodology provided by Matt Hansen of the University of Maryland (Hansen et al. 2008).

Given that clouds were then essentially “patched” in, this methodology increased the total number of forest and non-forest pixels both across the country and within the sampled cells. It must be noted that it was often very difficult to distinguish features properly due to the blurriness that the algorithm created. It's very important to note that there were some noticeable errors for some cells in the unsupervised classification for the gap-filled data that were due to the script not working perfectly. These were checked visually by MAE and Forest Carbon and corrected with various passes for some noticeably contentious areas.

Restratification and Revision of 2012 Sampling

Original 2012 Sampling Strategy[7]

As mentioned previously, MAE and the FAO worked earlier on a sampling approach to estimating forest cover for 2012 in order to save time and budget in achieving a total historical picture of forest cover in Ecuador from 1990 to 2012. In order to generate a sample population, MAE and FAO divided the country into 710 (19.8 km x 19.8 km) tiles stratified by total historical deforestation in each tile from 1990-2008. Each deforestation strata was randomly sampled (see Annex II: Equations Used to Calculate the Sample Size) based on the statistically required number of samples per strata. Stratified sampling is commonly used to subdivide sample population into relatively homogenous sub-populations, with lower within group variance. This in turn effectively reduces the size of the sample population (n_{strat}) necessary to reach a given confidence interval. In discussions with MAE it was agreed that it may be helpful to further expand the

[6] 75% overall accuracy of forest and non-forest classes is the lower cutoff threshold used by the Verified Carbon Standard’s Jurisdictional and Nested REDD (JNR) Requirements and was set as the classification threshold as part of this work.
[7] For a full explanation and analysis of this approach, see Eickhoff et al. 2013.

stratification from 3 layers to 4 or 5 and to base the threshold amount of deforestation between each of the strata on a statistically derived thresholds from natural breaks in the distribution[8].

The number of 2012 tiles needing to be sampled was also a function of the applied confidence interval (CI) and standard error (SE). In the original calculations done together with the FAO, MAE aimed for a 95% CI and 5% SE, a highly rigorous target[9].

One of the most important considerations for restratifying the deforestation layers was due to the total cumulative cloud coverage in each sampled tile and the Landsat 7 scanner line correction error much of the sampled area had no data thereby reducing the effective area sampled below the minimum intensity necessary for to reach the aforementioned CI. The original sample intensity based upon the three strata is displayed in the chart below.

Table 6: Original calculation of required number of stratified samples based on 95% Confidence Interval and 5% Standard Error, and corresponding effective sample sizes.

Strata[10]	Population (N)	Minimum Required Sample (n_h)	Actual Sample n_h – cloud>25%	Effective area sampled (cloud cover, no data, scanner line error)	Effective Sample
High (>20)	92	55	50	73%	36.33
Med (5-20)	195	100	86	68%	58.16
Low (0<5)	423	67	56	75%	40.35
Total	710	**222**	191		**137**

As can be seen, the total number of effectively sampled tiles was 137 as compared to the minimum 222 originally needed. Thus, there were essentially three options available to improve the sampling design:

1) reducing the size of the confidence interval,
2) the restratification of the sample population into more homogenous sub-populations and
3) increasing the sample size through the classification of ancillary images.

Of these options, changing the confidence interval was the simplest and cheapest option. However, due to the fact that the estimate of forest cover would become an official government figure and would be used to justify funding from the REM programme it was decided that the stricter confidence would be maintained. It was decided that restratification as well as increased coverage would be used to create a sample population that greatly exceeded the number of samples necessary to meet the confidence interval.

Restratification

A restratification was undertaken for 3, 4 and 5 strata classes using the JENKS Natural Breaks to determine thresholds for each and with the goals of reducing the minimum sample population necessary to reach the required CI and to improve the overall estimate of forest cover.

The equations used to reach these figures as well as associated spreadsheets can be found in Annex II and the overall sample population necessary (n_{strat}) as well as the strata sample populations (n_h) necessary to reach the desired CI and error rates.

[8] Method for determination used the JENKS Natural Breaks tool available in ArcMap.
[9] Guidance from the VCS and other common internationally recognized REDD+ thresholds typically default to a 90% CI/5% SE or a 95% CI/10% SE.
[10] Strata defined in terms of % Deforestation over the 1990-2008 period

Table 7: Summary of required number of samples required with 3 levels of deforestation strata

Original Stratification (percent deforestation 1990-2008)	N	Mean	SD	CV	Sum(% defor)	V	(CV*Vf)	nh
Low (<5%)	423	1.21	1.40	1.16	510	0.097	0.113	64.8
Medium (5-20%)	195	11.42	4.53	0.40	2227	0.424	0.168	96.8
High (>20%)	92	27.31	5.32	0.19	2513	0.479	0.093	53.6
	710				**5250**			
						wCV	0.37	
						t-value	1.96	
						nstrat	**215.26**	

A number of different restratification techniques were considered including using natural breaks, Jenks and Geometric Intervals to refine strata ranges. Ultimately, however it was decided to re-stratify the existing sample population using 5 strata. Strata were assigned using visually interpreted thresholds (in order to aid interpretation) at the intervals shown in the chart below. The restratification successfully reduced the overall strata size from 222 to 82 cells as seen by the chart below without compromising either the CI or SE.

Table 8: Summary of required number of samples required with 5 levels of deforestation strata

5 Strata (percent deforestation 1990-2008)	N	Mean	SD	CV	Sum(% defor)	V	(CV*Vf)	nh
Very Low (<1)	247	0.23	0.29	1.25	56	0.011	0.013	4.8
Low (1-5)	176	2.58	1.17	0.45	454	0.086	0.039	13.9
Medium (5-10)	90	7.28	1.44	0.20	656	0.125	0.025	8.7
High (10-20)	105	14.97	3.01	0.20	1572	0.299	0.060	21.3
Very High (>20)	92	27.31	5.32	0.19	2513	0.479	0.093	33.0
	710				**5250**			
						wCV	0.23	
						t-value	1.96	
						nstrat	**81.67**	

Results and Discussion

Change Analysis

Estimate of Deforestation

The deforestation rate was estimated by summing weighted strata means taken from the sample population defined by the stratification (the equations used for this calculation can be found in Annex III: Equations Used to Estimate the Rate of Deforestation). Based on the 5 strata sample used the gross annual deforestation rate in Ecuador between 2008-2012 was estimated to be 29,034.15 ± 4,604.46 ha. This is equivalent to a deforestation rate of 0.12% per year. The net annual deforestation rate was estimated to be 19,378.10 ± 4,579.46 ha or 0.06% per year. As can be seen in Table 9 below this represents an immensely significant drop in annual deforestation between the 2000-2008 and 2008-2012 accounting periods.

Table 9: Comparison of annual deforestation (ha) for gap-filled data

Deforestation Period	Gross Rate (ha/yr)	Net Rate (ha/yr)
2000-2008	115,075	75,287
2008-2012	29,034	19,378

To some extent the overall scale of this change is exaggerated by the low background deforestation rate. As any rate approaches zero the proportion of change is increased. We must also remember that deforestation rates across Ecuador were halved previously between the 1990-2000 and 2000-2008 accounting periods. This precipitous drop in deforestation rates could be attributed to a number of factors including the stabilization of agricultural frontiers, declining rural birthrates and increasing urbanization of the Amazon region (Carr, Pan, and Bilsborrow 2006) (Barbieri and Carr 2005). With that being said, a 75% drop in gross deforestation calls into question either the accuracy of the change assessment or the realistic accuracy of the gap-filled multi-temporal mosaic dataset.

Validation of the Sampling Methodology

Applying the Sampling Stratification to the Previous Time Periods

In order to verify the precision of the results produced by the stratification and the associated weighted deforestation rates described in the sections above, it was decided to compare the results of the estimator generated from the 2000-2008 data against the known deforestation rate from the wall-to-wall classification. The stratification used to estimate the deforestation rate for the 2008-2012 was used to select a sample population from the 2000-2008 dataset.

Table 10: Annual deforestation between 2000-2008

Change	Estimated (ha/yr)	Actual (ha/yr)	Percent difference (est – act)
Gross deforestation	109,274	115,076	-5.30%
Net deforestation	73,603	75,287	-2.30%

The estimator generated from the 2000-2008 data was shown to accurately predict the known deforestation rate (results in Annex III). This verified that the estimator reliably predicted deforestation rates based on a limited sample size. It should be noted that the use of 5-strata as opposed to 3 and 4 respectively neither significantly changed the estimated rate of deforestation nor the calculated confidence interval. The 5 strata approach did, however, significantly reduce the sampling intensity necessary to create a reliable estimator. This approach allowed for the removal of cells that were heavily obscured by cloud cover or had other problems related to visibility and/or classification errors.

Comparison of results with the original classification before the gap-filling

Having been surprised by the dramatic reduction in deforestation estimated between the years 2008 and 2012, an accuracy assessment and restratification to calculate gross and net changes were also conducted on the original Landsat imagery classification for 2012 without the gap-filling process. The following are the results that were obtained for the accuracy assessment of the non-gap-filled data, using the same RapidEye imagery as a reference.

Table 11: Accuracy assessment for the year 2012

Confusion Matrix: 2012 (without gap-filling)		**RapidEye**			*Classified (Landsat) pixels*	*User's Accuracy (%)*
		Forest	**Non-forest**	**Water**		
Landsat classification:	**Forest**	**200**	2	0	202	99
	Non-forest	26	**143**	1	170	84.1
	Water	0	**3**	**9**	12	75
Ground truth (RapidEye) Pixels		226	148	10	**384**	
Producer's Accuracy (%)		88.5	96.7	90		**91.6**

Overall accuracy = 91.6%
Average Accuracy = 91.73 %
K = 83.94 %

Using this classification and the equivalent 5 strata sample, the net annual deforestation rate was estimated to be 64,506 ± 22,790.44 or 0.26% per year over the same period. This would still represent a net reduction in annual deforestation but to a lesser degree than as compared to the gap filled dataset. A simple average between the two would not suffice as a resultant rate. This forces us to proceed with caution when assessing the validity of the gap-filled result.

However, the comparison of gross and net rates of deforestation between the final 'gap-filled' and initial 'non-gap-filled' analysis was not possible because the methods used were different and the input data was improved between the two analyses (Annex IV: Land Cover Classification).

Firstly, the initial analysis was not designed to produce gross rates of change, but rather net rates. The initial analysis used only a single data image segmentation and classification method to produce an independent assessment of forest area for 2012.

It is not advisable to compare two independently produced maps to estimate gross changes between time periods as any errors in the classification of land cover is compounded during the estimation of changes. The final 'gap-filled' analysis used a multi-date segmentation which produced only a single polygon layer incorporating spectral information from both 2008 and 2012 time periods. The polygons could be labeled twice, once for 2008 and again for 2012, but any changes in land surface should have been more accurately captured in the polygon layer enabling year-on-year tracking of a single point on the ground; and, thus, more accurate estimates of gross deforestation.

Conclusions and Recommendations

Sampling Methodology

The MAE/FAO sampling methodology has been tested and under the new stratification and sampling intensity has been shown to be able to deliver results within 2.3% of actual calculated net values and 5.3% of gross values while sampling more than 60% fewer plots and still achieving a 95% CI and 5% SE. This is quite accurate, with an error margin of only approximately 1,680 hectares across the entire country over the 2000-2008 period.

Sources of Potential Error

Valid measures of map accuracy are critical, yet can be inaccurate even when following well-established procedures. Accuracy assessment is particularly problematic when thematic classes lie along a land-cover continuum, and boundaries between classes are ambiguous.

In contrast, the current accuracy assessment was based on higher-resolution imagery and only considering a simpler differentiation of forest to non-forest, with cloud and band gaps filled as well as further visual verification and correction.

However, it must be noted that due to machine processing limitations, it was only possible to run a *simple* random sampling on the 2008-2012 cell sample. It would have been more accurate to run a *stratified* random sampling based on coverage of the land use changes, with priority weighting on deforestation and regeneration. It might be worthwhile running one for extra assurance based on a systematic selection of, for example, each 30 segments classified as deforestation and each of regeneration.

represents the landscape (forest and soil) on the left. On the right however, is a land cover map of this area with green representing forest. The entire area would be mapped as forest.

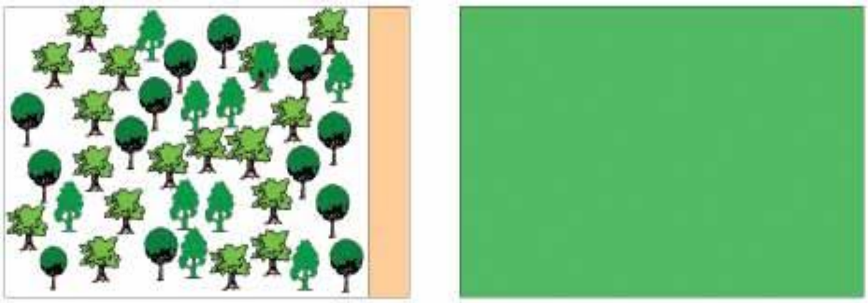

Figure 7: Generalization in remotely sensed land use classifications

In this figure we can see that the classification did not accurately represent the actual landscape since it classified the entire area as forest whereas 10% is actually non-forest. In accuracy assessments the overall map accuracy would be 90%, but one could say that the accuracy of the forest class is 100% since all of the entire area of the forest class was accurately classified as forest.

This is an exaggeration, but we must bear in mind that forest cover from low resolution satellite imagery is essentially an estimate, especially when applying multi-temporal imagery to compensate for gaps in the data; and more so in areas with seasonal variance (for example in Loja and Manabi on the coast of Ecuador).

We must also bear in mind that the land classified as forest followed the definition from the Marrakech Accords (UNFCCC, 2001). Under this agreement forest is defined as: a minimum area of land of 0.05-1.0

hectares (ha) with tree crown cover (or equivalent stocking level) of more than 10-30% with trees with the potential to reach a minimum height of 2-5 m at maturity in-situ.

This amount of crown cover in a tropical rainforest environment is actually relatively low, so we must also proceed with caution when making assumptions about highly fragmented landscapes where the extent and impact of fragmentation is a lot of the time impossible to measure before it can truly be considered deforestation under this definition. Higher-resolution imagery would be necessary in these suspected areas to realistically measure and monitor deforestation.

Historical Mapping

We find that both the gap-filled and non-gap-filled historical imagery provides accuracy assessments that are not only higher than the minimum required by the VCS JNR requirements (75%) but also higher than are typically seen on many provincial level classifications.

Upon closer inspection the gap-filled historical imagery may not entirely represent the landscape dynamics due to the blurriness of the imagery, especially in gap-filled areas (Figure 8, Annex IV: Land Cover Classification). However, the historic land cover map of Ecuador for 2008 was improved between the two analyses, so comparing the results generated from the initial with those of the final analysis then is thus impossible.

At best, we can say that the data is uncertain. In situations of uncertainty it is wiser to reduce risk and exercise caution and conservativeness. We recommend that until a complete high-resolution 2013 wall-to-wall map is available, the MAE should exercise caution and recognize that we've confirmed a reduction in deforestation, but are still unsure of the exact amount and rate.

Once the 2013 wall-to-wall analysis has been completed, the quantities will be more easily rectified and consolidated against the 2013 results.

If the MAE follows the gap-filled results as fixed quantities, rather than as an estimated tendency, it may open the Ministry up to criticisms of using a flawed or imprecise multi-temporal data model to establish a historical baseline and may inadvertently put itself in the position where the 2013 data shows that the reality is actually a less dramatic decrease in deforestation.

By taking this approach in the short-term, a mid-point result from the 2013 data could result in a further decrease in deforestation, as compared to an increase.

Other Recommendations and Suggestions

Further validation

Further validation testing before 2013 could include trying to use just the proportion of forest in each cell to estimate the total forest area for the whole country, the same way that it was done for the change calculations, except using the proportion of forest instead of the proportion of change. This could help to understand better where any underestimations may have taken place. Theoretically, the forest area for 2012 ought to be similar. Another validation of accuracy could be to apply the sample wall-to-wall segmentation to the earlier period in order to validate the precision of that rate calculation.

Software suggestions and capacity building for future efforts

If the MAE will need to use legal software in the coming year, then if they manage to acquire ArcGIS 10+ licenses, they will profit significantly from learning how to apply batch processing with ArcGIS 10+'s ModelBuilder Iteration functionality. Until now they have been going through many of their vector and attribute geoprocessing tasks very manually. Having to select every single combination and calculating on these individually for every single layer. When we're dealing with the processing for a whole country, this is truly inefficient and prone to error. Together with work flow modeling, they would also benefit from basic programming skills in Python, which would be useful not only for ArcGIS, but also for the free, open source alternative of QGIS.

Otherwise they will need to consider how to run iterative models with the Modeler now available in QGIS through the Sextante toolkit. From here, image segmentation would also be easily possible with Montedverdi's Orfeo Toolbox plug-in, as well as supervised and unsupervised classifications. A more robust and dedicated raster processing software would be Spring, however this has a much steeper learning curve. Demonstrations were provided during the consultancy and some of the staff had already been interested or exposed to these alternatives.

Annex I: Forest Definition

Forest cover changes were based on the Level 1 classification schema used by the MAE is shown in Table 12 below. This conforms to the six broad land use categories in accordance with Intergovernmental Panel on Climate Change (IPCC) reporting guidelines.

Table 12: MAE Level 1 Classification System

Classification	Definition
Forest	Plant community of at least one hectare with trees of 5 m in height and with a minimum of 30% canopy cover. Includes: areas with bamboo and native palms, provided they meet the minimum limit for height and canopy cover. Excludes: tree stands in agricultural production systems, for example in fruit plantations, oil palm plantations and agroforestry systems. Also excludes trees growing in urban parks and gardens.
Shrub and Herbaceous Vegetation	Areas covered by shrubs and herbaceous vegetation, from natural biological processes, not including agricultural areas.
Agriculture & Livestock	Planted grass and areas under agricultural cultivation, or areas that are within a rotation between the two.
Water bodies	Area that is covered or saturated by static or moving water, natural or artificial, resting on the earth's surface for all or part of the year.
Urban zones	Human settlement and infrastructure that complements it.
Other lands	Areas with little or no vegetation, rocky outcrops, glaciers and other classes that are not included in any of the other categories.
No Information	Applies to areas that have not been mapped.

Annex II: Equations Used to Calculate the Sample Size

The following equations were used in the calculation of the total number of required tiles to be sampled per strata layer.

$$n(st) = \frac{t^2 * (wCV)^2}{A^2}$$

Eq. 1: The number total sample size was calculated using the following equation

Where:

A	= percent admissible error (5%)
n(st)	= sample size
t	= t-value based on desired confidence interval (95%)
wCV	= the weighted coefficient of variance calculated for all strata calculated

$$wCV = \sum (vf_h * CV_h)$$

Eq. 2: The weighted coefficient of variance is calculated using the following equation

Where:

$$CV = \frac{\sigma}{|\bar{x}|}$$

Eq. 3: Equation for the calculation of the coefficient of variation of the strata

Where:

σ = the within strata standard deviation (percent deforestation 1990-2008),

x = the within strata mean

vf_h = the proportion of deforestation represented by the strata within the total population.

$$n_h = \frac{n * (N_h * \sigma_h)}{\left[\sum (N_i * \sigma_i)\right]}$$

Eq. 4: Neyman Equation for calculation of the number of samples required per strata.

Where:

n_h = The required within strata sample size,

n = The total sample size

N_h=The total strata size

Annex III: Equations Used to Estimate the Rate of Deforestation

The following equations can be followed in spreadsheet format Appendix II.

$$D_a = \frac{a_c * W\mu}{t}$$

Eq. 5: Equation for the estiamtion of deforestation as an extrapolation of a sample

Where:

D_t = Estimated deforestation rate (ha^{-1} yr^{-1})

a_c = the area of the Ecuador (24,974,309 ha)

W_μ = the weighted strata mean (% deforestation)

t = the number of years within the given calculation period

$$W_{\mu} = \sum(\mu_i * m_i)$$

Eq. 6: Equation defining the weighted strata mean.

Where:

μ_i = the mean rate of deforestation within a given strata

m_i = the stratum weight

$$m_i = \frac{N_i}{N_p}$$

Eq. 7: Equation for defining specific strataum weight.

Where:

N_i = The population of a given strata (cells)

N_p = The total population (710 cells)

Forest Carbon
Mitigating Climate Change Through
Conservation & Sustainable Forestry

Annex IV: Land Cover Classification

The improved Landsat mosaics for years 2008 and 2012 were subset according to the original 222 selected sample sites. At each site, Landsat bands 4, 5, and 7 were combined in a multi-date stack. Image segmentation was applied to the multi-date stack to produce a single set of image segments suitable for land cover classification in each time period and changes between time periods[11]. The multi-date nature of the segmentation allows for calculation of gross and net changes and reduces errors of omission and commission that may occur in individual map comparisons.

Image segments were labeled first with land cover classifications corresponding to the improved, historic 2008 land cover map of Ecuador following a simple majority rule. An automated change detection process was applied to the Landsat imagery to determine those segments most likely to have experienced a change in land cover between 2008 and 2012. Where segments were not detected as changed, the land cover label from 2008 was assigned directly to the segment in 2012. Segments indicating a change in land cover were labeled using a supervised classification approach with training data derived from surrounding, non-changed segments[12]

It must be noted that, due to remaining radiometric differences between original and 'gap-filled' pixels (caused by the absence of any high quality pixels, seasonality differences, etc.), image interpretation and feature delineation remained very difficult; the blurriness of which can be seen in Figure 8. It is for these reasons that the processing chain included the critical step of reviewing and revising the results of the automated classification in order to ensure as accurate a result as possible. Visual review and revision of the automated classification results was conducted by experts from the MAE, Forest Carbon and FAO.

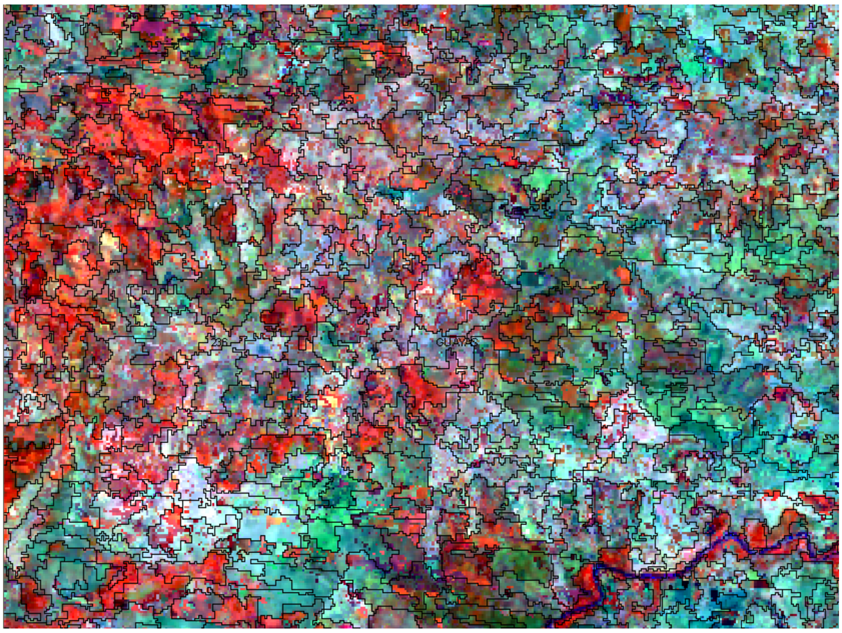

Figure 8: Cell 236, example of blurriness in gap-filled cloudy areas.

[11] Duveiller, G., Defourny, P., Desclee, B., & Mayaux, P. (2008). Deforestation in Central Africa: Estimates at regional, national and landscape levels by advanced processing of systematically-distributed Landsat extracts. *Remote Sensing of Environment*, *112*(5), 1969–1981. doi:10.1016/j.rse.2007.07.026

[12] Lindquist, E. J., D'Annunzio, R., Gerrand, A., MacDicken, K., Achard, F., Beuchle, R., … Stibig, H.-J. (2012). *FAO Forestry Paper 169: Global forest land-use change from 1990 - 2005* (p. 40). Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/docrep/017/i3110e/i3110e.pdf