Constructing regional baselines for carbon emissions from land use change in Chiapas, Mexico

A. Hellier[1], M.A.Castillo-Santiago[2], R.Tipper[2] and B.H.J. de Jong[3]

---

[1] Greenergy Carbon Partners & ECCM; 18, Liberton Brae; Edinburgh EH16 6AE; Scotland richard.tipper@eccm.uk.com

[2] El Colegio de la Frontera Sur (ECOSUR); Unidad San Cristóbal de las Casas, A.P. 63; C.P. 29200; San Cristóbal de las Casas; Chiapas, México; e-mail: mcastill@sclc.ecosur.mx

[3] ECOSUR; Unidad Villahermosa; A.P. 1042, C.P. 86100, Col. Atasta, Villahermosa, Tabasco, México. E-mail: bjong@vhs.ecosur.mx

Abstract

We present a methodology for constructing and testing regional baselines for carbon emissions from land-use change using a spatial modelling approach for an area of 2.7 million ha in southern Mexico. The methodology is based on an analysis of causal factors of previous land-use change (Castillo *et al.* 2002). Carbon risk models using the causal factors identified were used to calculate expected future carbon emissions. The performance of each model was tested by estimating carbon emissions 1975-1996 from randomly selected sample plots in the study area and comparing the results with observed carbon losses over the same period. Errors were relatively low: less than 5% of vulnerable carbon stocks for areas in excess of 10,000ha and around 10% at a scale of 2,500ha. The methodology provides an objective means of constructing baseline scenarios and setting risk buffers using readily available cartographic and census data that could be used to assess the potential for forest conservation projects in developing countries to offset carbon emissions.

Key words: baselines, carbon emissions, deforestation, forest conservation, Mexico, risk buffers

1 Introduction

Assessments of the greenhouse gas benefits of forest conservation projects requires the construction of so-called “baseline scenarios” that describe the expected status of the terrestrial carbon stocks in the absence of the project. However, no standard methods currently exist and recent pilot projects have used a number of quite different approaches:

- Extrapolation of past trends – e.g. the Norway-Costa Rica AIJ project in the upper Virilla river basin. The baseline assumed that a local deforestation rate of 7.5% 1986-1992 would continue for the next 20 years (UNFCCC 2000).
- Hypothetical future scenarios – e.g. the Rio Bravo Conservation Management Area, Belize. The baseline is defined by the intent of key stakeholders to purchase the land for conversion to agriculture (Stuart and Moura Costa 1998).
- Prevailing technology or practice – e.g. the ICSB-NEP reduced impact logging project in Malaysia. The baseline assumes that current logging practises would continue without intervention (Stuart and Moura Costa 1998).

The perceived technical difficulty of setting baselines for forest conservation projects in developing countries in a consistent and comparable manner was one of the reasons why the United Nations Framework Convention on Climate Change (UNFCCC) did not include this type of activity in the Clean Development Mechanism for the first commitment period 2008-2012.

The existing methods often either fail to capture regional variation in the causes of carbon emissions or are not based on scientific and objective methodologies. None allow an objective assessment of whether the baseline is appropriate to the area in question or provide a measure of how accurate the prediction is likely to be. Extrapolation of past trends can be supported by scientific evidence but takes no account of the regional variation in the processes that cause deforestation. Hypothetical future scenarios can take account of local details in land-use patterns but are hard to standardise and could be abused by those seeking to over-state project benefits. The assumption that current practises will continue into the future does not take into account political and financial pressures to improve management practises (e.g. low impact logging) that could also reduce carbon emissions.

To produce credible emissions reduction units through the conservation and management of forests, verifiable, evidence-based, standardized methodologies are required to set baselines. As there is often significant variation in the socio-economic conditions within any region the methodology should take into account regional trends in land-use and local differences in the way that rural communities manage their resources. An objective means of assessing the accuracy of a proposed baseline is also required so that estimates of emission reductions may be modified to include a risk buffer (carbon that should not be sold or used to offset other emissions).

In this study we illustrate the application and testing of a spatial modelling approach to construct a multi-project baseline for an area of 2.7 million ha in southern Mexico.

2 Methods

The analysis presented here employs data from a study of the causal factors of land-use change and carbon emissions over a period of 21 years between 1975 and 1996 for an area of 2.7 million ha in the north-east of Chiapas, southern Mexico (Castillo *et al.* 2002). The study area included three environmental and demographically defined regions: the Highlands, the mid-elevation Cañadas and the Selva lowlands and identified three socio-economic factors as being closely related to land-use change (see table 1). A detailed description of the study area and the relationships between selected factors and land-use change are given in the accompanying paper. In this paper we use the results generated by Castillo *et al.* to calculate baselines emission for forest conservation projects and associated risk buffers, taking into account local conditions, current land-use, and the values of the predisposing and driving factors for specific test sites.

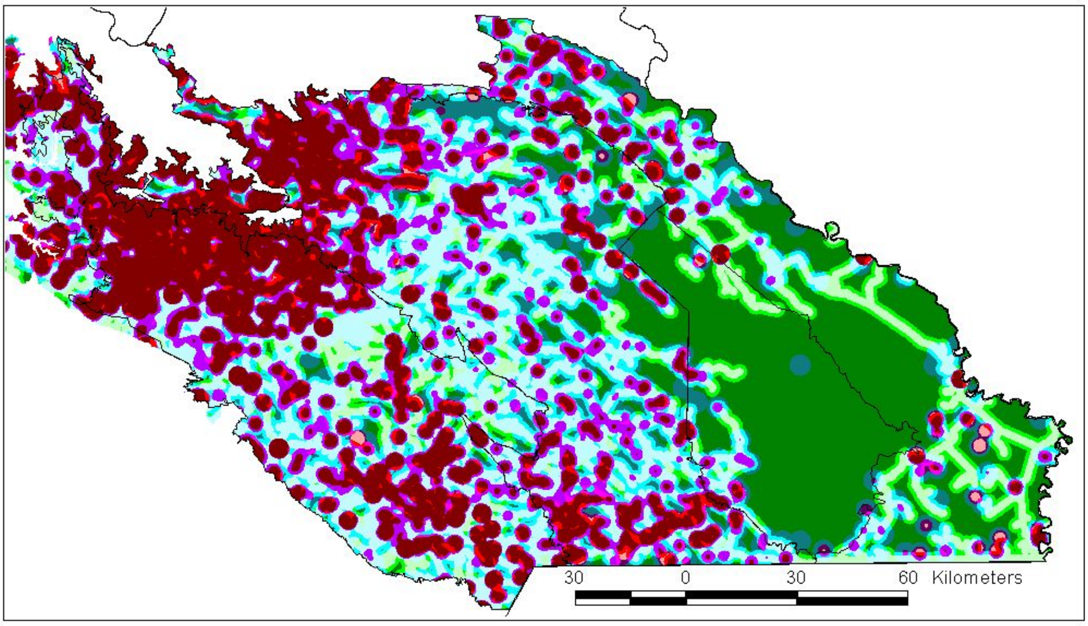
Table 1. Causal factors of land-use change applied in the analysis

Causal factor	**Definition**	**Classification**	**Source**
distance of forest from roads	distance from paved and unpaved roads built up to 1980	0 to1000, 1000 to 2000, more than 2000m	1:50,000 road maps (citas)
distance of forest from agriculture	distance from agriculture, pasture and disturbed land; in 1975	0 to 500, 500 to 1000, more than 1000m	1:250,000 INEGI land use maps (citas)
Farmer density	population whose primary occupation was farming in 1990	0, 0 to 15, 15 to 30, more than 30 /km2	INEGI 1990 population census (cita)

Carbon risk models in the form of matrices using the causal factors of land-use change identified in Castillo *et al.* (2002) were used to calculate expected carbon emissions. Each matrix combined a predisposing and a driving factor to give 12 categories of risk of carbon emissions (represented by the 12 cells in the matrix). Two models were tested: distance to roads combined with population density (hereafter called DistRd); and distance from agriculture combined with population density (DistAg). For each category in the matrix, the historical loss of vulnerable carbon between 1975 and 1996 was used as the input value. Each model was parameterised with data derived from three spatial scales: the whole 2.7 million ha study area, the three sub-regions (Highlands, Cañadas and Selva) and a case study area within the Selva lowlands: Marques de Comillas (comprising of two municipalities, approximately 200,000 ha in total).

Maps with the spatial distribution of the carbon risk categories of each model were created in a Geographic Information System (GIS). An example of a risk map for the DistRd model parameterised at the scale of the study area is given in figure 1 (a full list of models used in the analysis is given in Appendix 1).

Figure 1: Risk map for Distance to Roads model (parameterised with data from the 2.7Mha study area) showing percentage of vulnerable carbon lost over the period from 1975 to 1996



		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	52.1%	43.2%	27.5%
	>15 to 30	43.6%	36.6%	28.0%
	>0 to 15	36.8%	27.9%	20.9%
	0	32.1%	19.7%	5.8%

2.1 Testing model performance

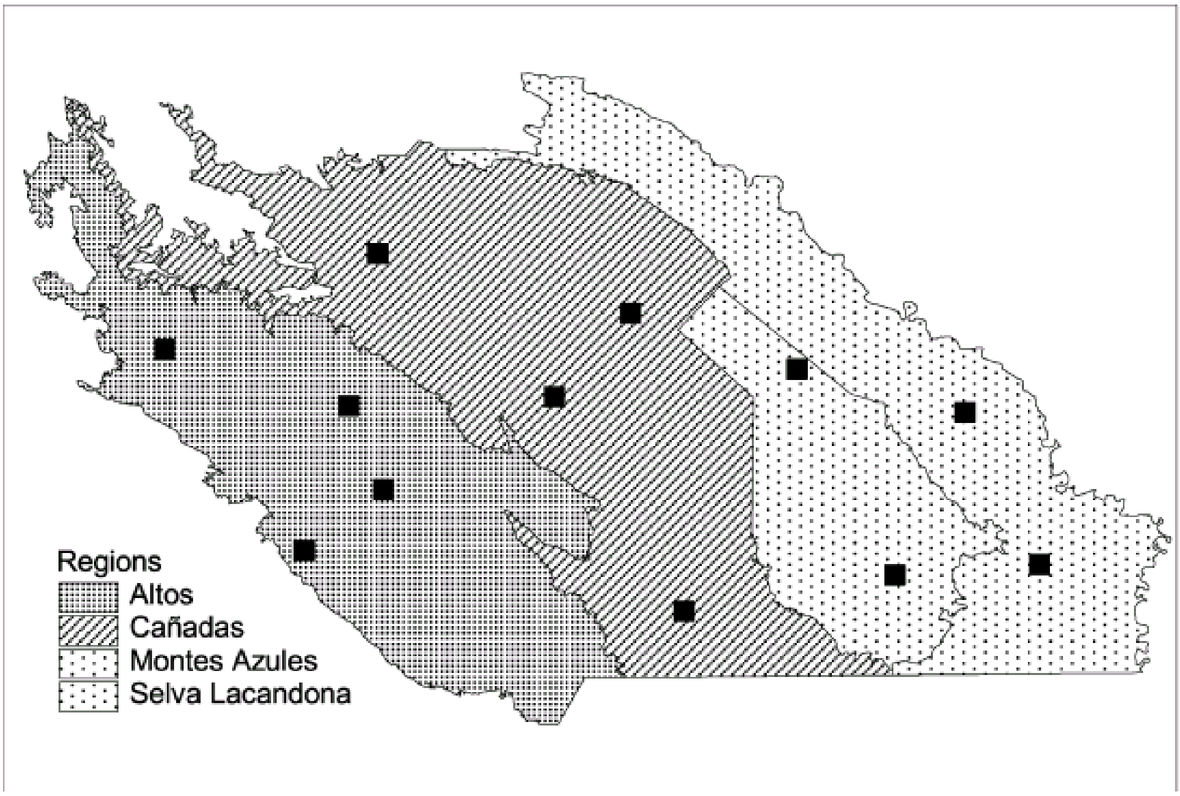
The performance of each model was tested by applying the risk models to fifteen 5 x 5 km (2,500 ha) randomly selected sample plots across the study area (Figure 2). The spatial distributions of the 12 risk categories of each matrix were mapped in each 2,500 ha sample unit. These maps were then intersected with the 1975 vegetation map and total expected carbon emissions for 1975 to 1996 were calculated for each category in the matrix. The results were then compared to observed carbon losses over the same period. The error for each model was expressed as: observed loss of carbon as a percentage of total vulnerable carbon stock minus expected loss of carbon as percentage of total vulnerable carbon stock.

Performance was assessed at three spatial scales:

- At the 2,500 ha scale: error was calculated by taking the mean of the root square of the error for each sample unit in a particular region.
- At the 10,000 ha scale: error was calculated from the total carbon loss within the combined areas of the four sample units in each region
- At the 37,500 ha scale: error was calculated from the total carbon loss within the combined areas of all 15 sample units across the whole study area (37,500 ha).

The size of the basic sample unit (2,500ha) was chosen to reflect the area of a typical community forest project. The error, therefore, gives an indication of what level of accuracy may be expected if the models are used to predict future carbon emissions at this spatial scale. (Note: the mean root square function for this scale is used so that an under-estimate in one sample unit does not counteract the effect of an over-estimate in another unit, error is therefore representative of a randomly selected 2,500 ha plot). Errors from the application of the models to the combined 10,000 ha area in each region and to the combined 37,500 ha across the whole study area give an indication of accuracy of predictions at the scale of a larger forest management or conservation project.

Figure 2: location of sample plots in study area



2.2 Application of the models to predict future emissions

To illustrate the application of this approach for project baseline construction, the risk models were used to predict emissions for the next 21 years (by overlaying a map of current vegetation and carbon density with the risk map generated from each model). In order to calculate the allowable amount of carbon credit accruing to a project we used the error calculated for each model to define the size of the risk buffer to allow for uncertainty in the baseline prediction at the relevant scale of application.

The site used to demonstrate the methodology was La Corona, a community of approximately 2,200 ha located in Marques de Comillas. Land-use data from 2000 were available for this community and predicted emissions from 2000-2021 were calculated. Total carbon emissions from the whole 2.7 million ha study area were also estimated, in this case using land use data from 1996 to give expected emissions up to 2017.

3 Results

3.1 Project scale model performance

Tables 2 to 4 show the results of the error analysis of the DistRd and DistAg models for the three sample sizes. Figure 3 shows the relationship between model error and sample size. When the sample size is at a project scale (2,500 ha), local variation in the relationship between carbon emissions and predisposing/driving factors has a significant effect on the model performance. However as the sample size increases it is more likely that the relationship observed in the sample will be the same as that observed for the whole region and the error decreases. Applying the model to the combined 37,500ha across the whole study area produces an error as low as 2.6% of total vulnerable carbon for the DistRd model and 6.9% for the DistAg model.

In most cases the error in estimated emissions was lower when the model was parameterised at the regional scale than when data from the whole study area were used (Tables 2 and 3). The relationship between carbon emissions and causal factors varies from region to region and it therefore is logical that models parameterised at the regional scale will be better predictors of emissions than models which use data from a larger area that encompasses a range of land use change patterns and ecological conditions. However, in the Selva region the errors related to the regional-level analysis were very similar to those observed at the level of the whole area. This is probably due to the variation in land use in the Selva, which contains both the relatively undisturbed Montes Azules Biosphere reserve as well as areas that have experienced high rates of land use change. Only when models were parameterised with data from the Marques de Comillas municipality were errors substantially reduced (see figure 4). This implies that for the Selva further sub-division of the region into distinct areas will be warranted (See also De Jong *et al.* 2000).

Table 2. Average error as a percentage of vulnerable carbon stock for 2,500ha sample units by matrix and parameterisation scale for each region

Model	**Scale of parameterization**	**Sub-region**			
		Highlands	**Cañadas**	**Selva**	**Marques**
DistRd	Whole area	18.7	12.0	5.5	12.9
	Sub-region	18.0	11.3	5.3	11.5
DistAg	Whole area	13.6	12.0	16.3	11.1
	Sub-region	13.7	12.7	16.2	7.3

Table 3. Error as a percentage of vulnerable carbon stock for combined 10,000ha sample area by model and parameterisation scale for each region

Model	**Scale of parameterization**	**Sub-region**			
		Highlands	**Cañadas**	**Selva**	**Marques**
DistRd	Whole area	4.2	7.3	3.1	3.7
	Sub-region	0.0	5.2	3.2	0.9
DistAg	Whole area	0.5	8.3	7.7	9.4
	Sub-region	1.9	5.5	8.6	0.2

Table 4: Error as a percentage of vulnerable carbon stock for combined 37,500ha sample area by model for study area

Model	**Error**
DistRd	2.6
DistAg	6.9

Figure 3: relationship between model error and sample size in each region for the Distance to Roads (DistRd) model, parameterized at two spatial scales

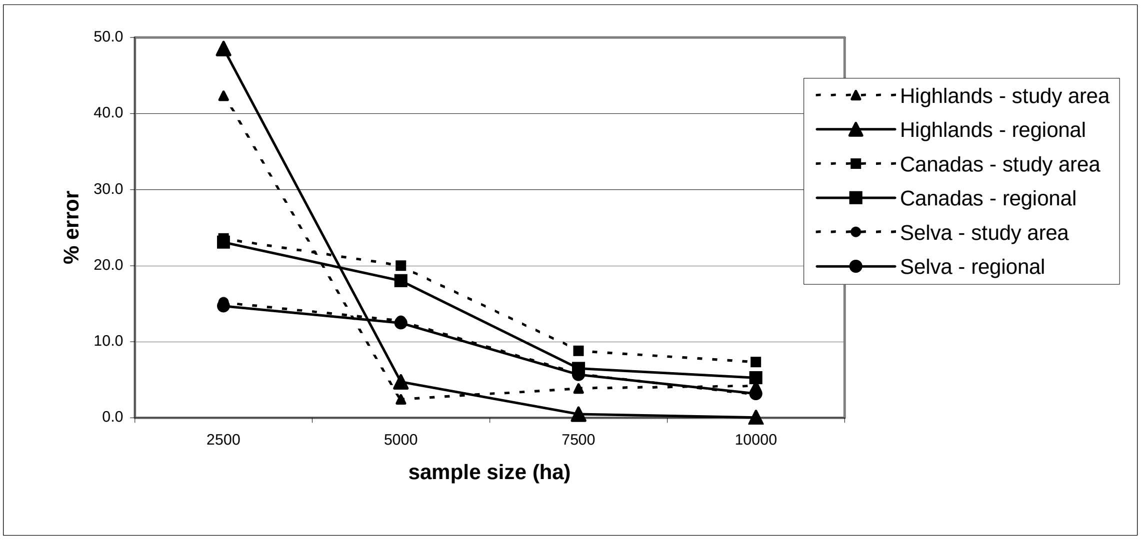
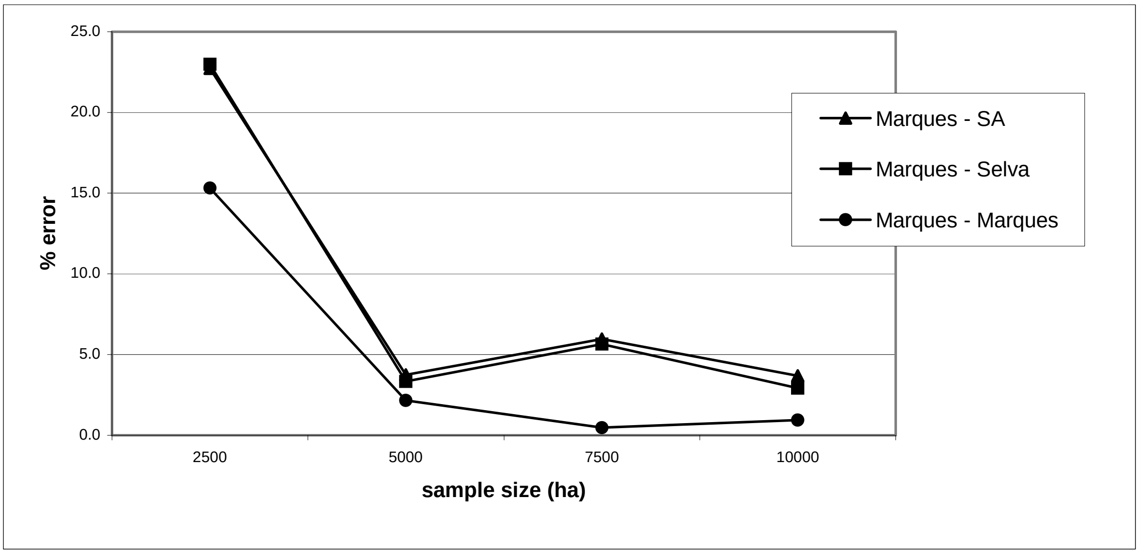


Figure 4: relationship between model error and sample size in Marques de Comillas for the Distance to Roads (DistRd) model, parameterized at three spatial scales



In the Highlands the DistAg model produced the lowest errors for the 2,500 ha sample units. While both models produced similar results at the 10,000 ha scale land use patterns in this region would suggest that the DistAg model is most appropriate. The population in the last 20 years has been relatively stable with little colonisation of new areas and deforestation has mainly occurred through the piecemeal expansion of agricultural lands. In the Selva the DistRd model consistently produced the lowest errors. Again this is consistent with land use patterns in this region, which had relatively little agricultural land at the start of the study period and has been subject to high levels of in-migration over the past 20 years facilitated in part by the road network. This result, however, was not observed in the Marques de Comillas municipalities where the DistAg model produced lower errors. This was not expected as demographic patterns and land-use change processes have been similar in Marques and the rest of the Selva region. While this could be due to differences in land-use change processes it is more likely to be a facet of the sample unit selection and for this reason the DistRd model is considered to be the most appropriate for the Marques region. In the Cañadas, both models produced similar results; in this region deforestation has occurred partly through the incremental expansion of agriculture but has also been driven by colonisation of new areas facilitated by the presence of roads. It would therefore be possible to use either model for the Cañadas but the DistRd model produces a slightly better performance.

3.2 Project level application – the community of La Corona, Marques de Comillas

Figure 5a shows land use in La Corona in the year 2000; the vulnerable carbon stock in each vegetation type is given in table 5. Land use is derived from ET-M Landsat satellite image of 2000 and vulnerable carbon stocks from data given in Castillo *et al.* (2002). Based on the above analysis of model performance the DistRd model parameterised with data from the Marques de Comillas municipalities was used to predict carbon emissions from this area. Figure 5b shows the spatial distribution of the risk classes of this model across the community land. The expected carbon emissions from La Corona are given in table 6; total expected emissions 2000-2021 from La Corona were calculated as 78,991 tC. In theory, therefore, if the forest resources of La Corona could be conserved 78,991 tC emissions could be avoided over the next 21 years.

Figures 5: Land use in La Corona and spatial distribution of Distance to Roads model risk classes in 2000

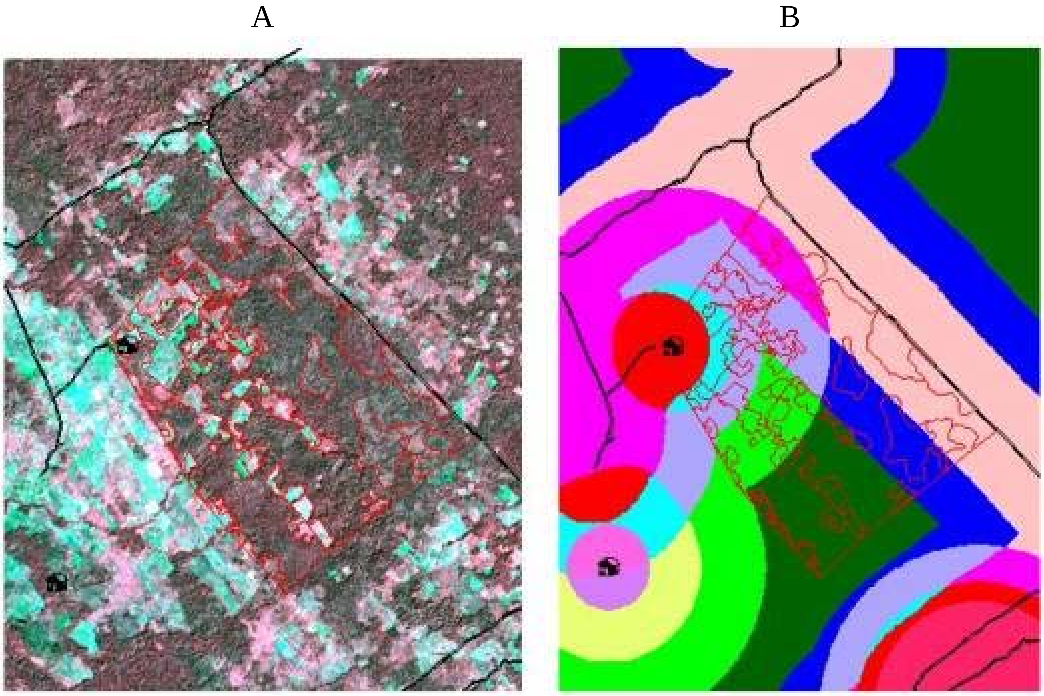


Table 5: Vulnerable carbon stock in La Corona by vegetation type in 2000

Vegetation type	**Area (ha)**	**Vulnerable carbon (t)**
Forest	1,238	273,940
Disturbed forest	332	34,477
Secondary vegetation	253	0
Agriculture	247	0
Pasture	188	0
Settlement	17	0
Total	**2274**	**308,417**

Table 6: Predicted emissions from he community of La Corona 2000-2021 (tC)

		Distance from roads (m)		
		<1000	**1000-2000**	**>2000**
Population density (/km2)	**>30**	0	0	0
	15-30	1,802	1,409	0
	0-15	7,034	15,122	8,829
	0	21,311	10,467	13,017

In order to assess the amount of carbon that could be offset through such activities a risk buffer must be set to account for possible inaccuracies in this baseline prediction. The community of La Corona covers 2,274ha, approximately equal in size to the sample units used in the analysis. The error in the DistRd model for the Marques region was 11.5% of vulnerable carbon stock at the 2,500 ha scale (table 2). For the La Corona community this equals to a risk buffer of 35,468 tC. The predicted offset potential of conserving forest resources in La Corona 2000-2021 is therefore 78,991 - 35,468 = 43,523 tC.

3.3 Regional predictions of future carbon emissions

Table 7 shows predicted carbon emissions from the study area and each region over the 21 years from 1996 to 2017. Based on the analysis of model performance the DistRd model was used to predict carbon emissions from the study area and in the Cañadas and Selva regions and the DistAg model for the Highlands. The range of estimated emissions given in table 7 is based on the errors produced by the application of these models at the 37,500 ha scale for the whole area and the 10,000 ha scale for the separate regions. Based on these values (given in tables 3 and 4) conservative error ranges of 5% vulnerable carbon for the study area, Highlands and Selva regions and 10% for the Cañadas were used. Figure 6 shows the distribution of predicted emissions

1996-2017 across the study area. Figure 7 shows the observed change in vulnerable carbon stocks in the study area between 1975 and 1996 and the predicted change between 1996 and 2017 using the DistRd model with the 5% risk buffer.

Table 7: predicted carbon emissions 1996-2017

Region	**Model (and scale of parameterisation)**	**Vulnerable carbon 1996 (MtC)**	**Predicted emission 1996-2017 (MtC)**
Study area	DistRd (study area)	273.33	67.17 ± 13.7
Highlands	DistAg (regional)	36.74	11.21 ± 1.8
Canadas	DistRd (regional)	78.65	6.06 ± 7.9
Selva	DistRd (regional)	146.98	23.79 ± 7.3

Figure 6: Distribution of predicted carbon emissions from the study area 1996-2017 using the DistRd model

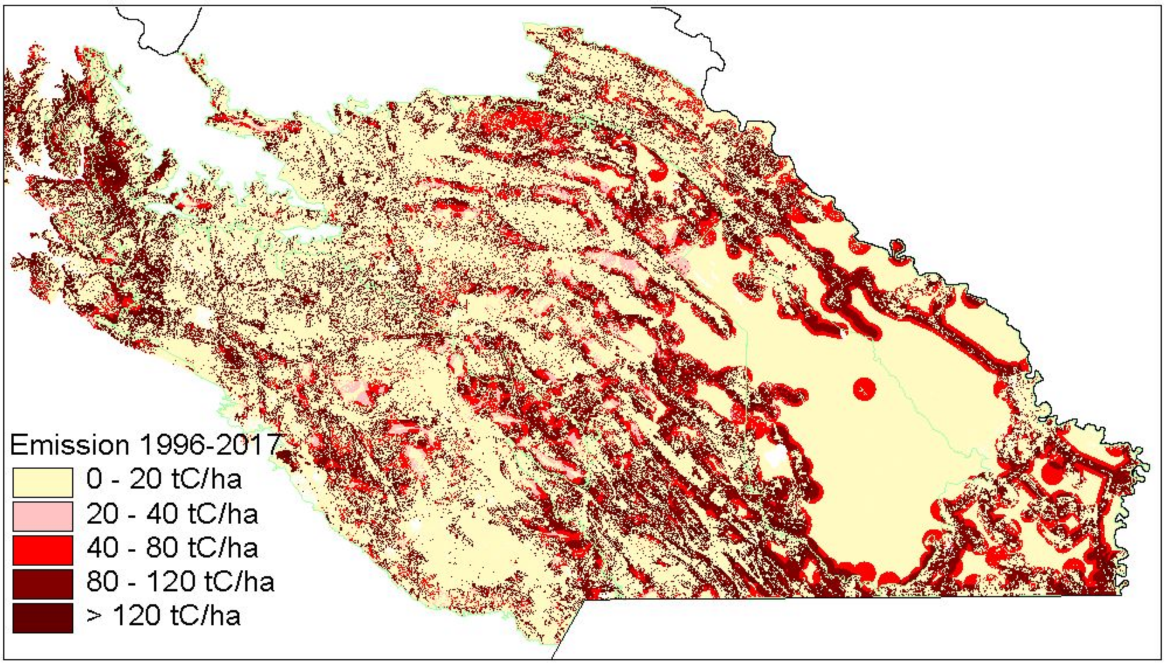
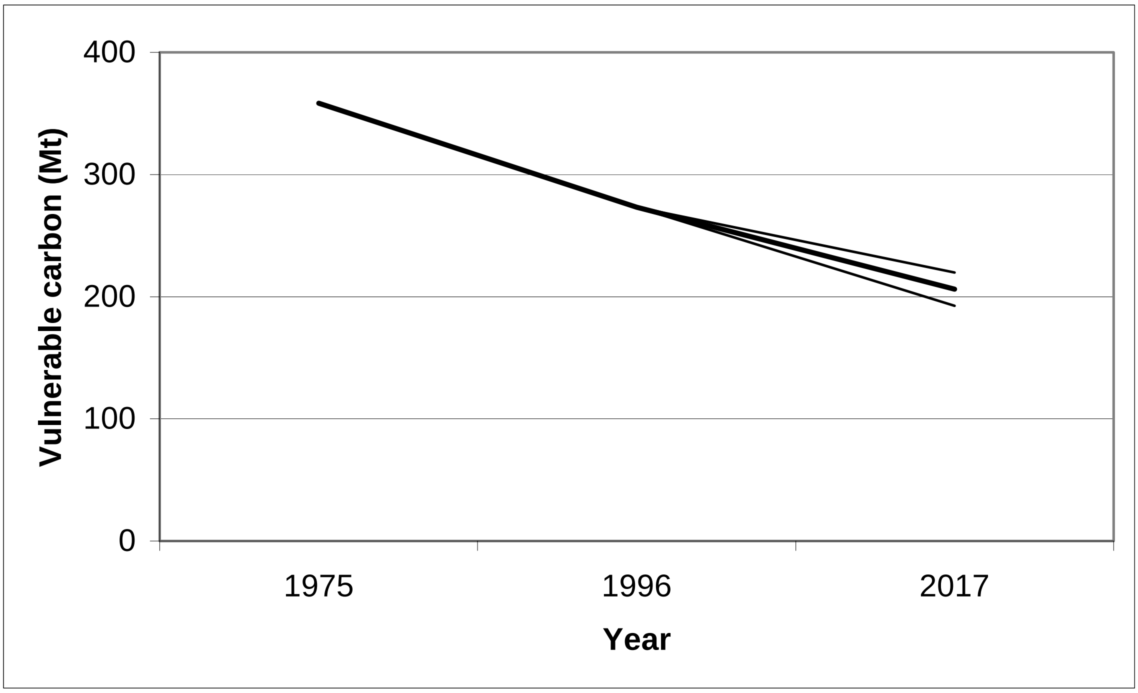


Figure 7: Predicted change in vulnerable carbon stock in the whole area between 1975 and 2017 using the DistRd model including the upper and lower limits of the risk buffer.



4 Discussion

There is considerable potential to conserve carbon stocks through conservation and sustainable forest management activities in Chiapas. Predicted carbon emissions from land use change from 1996 to 2017 for the 2.7 million ha study area are 67 MtC, equal to approximately half total UK annual emissions. While not all these emissions may be avoidable, forest management and conservation activities in this area could have a high potential to mitigate carbon emissions. This study has produced a methodology for constructing evidence-based regional baseline scenarios through an analysis of the relationship between land-use change and prevailing socio-economic conditions, using readily available cartographic and census data. The methodology provides an objective means of selecting the most appropriate baseline for an area through an analysis of the error produced when the models are applied to sample areas. This gives an indication of what level of accuracy can be expected when these models are used to predict future carbon losses and can be used to assess model performance and to set a suitable carbon risk buffer.

While the ideal methodology for assessing model performance would involve the use of data from a different period to that used in the initial analysis this was not possible here because of the scarcity of available data. However, scarcity of data is a problem encountered in many developing countries and the methodology described here provides an alternative means of testing model performance, which will be more widely applicable than a method that relies on very large datasets.

It should be noted that while this assessment gives an indication of the accuracy of predicting future carbon emissions this will also depend on the extent that the relationship between deforestation and the predisposing/driving factors observed in the past remain the same for the next 20 years. In areas such as the Highlands where there is little new colonisation occurring it may be reasonable to assume that the relationship between deforestation and the causal factors used in the analysis will be similar in the future. However, in other areas where there have been marked changes in population dynamics in the last two decades, it is likely that recently observed relationships and the spatial variation in these relationships will change in future. In the Cañadas, for example, where the DistRd model produced the best performance when tested on 1975-1996 data it is possible that greater population stability in this region will mean that DistAg would be a better predictor of land use change in the future. The selection of a model to predict carbon emissions should therefore not solely be based on the results of model performance but should consider likely changes in land-use patterns in the future.

By providing objective means of constructing baseline scenarios and setting risk buffers based on evidence of causal factors of land use change this methodology answers a number of the criticisms of current methods for calculating the carbon benefits of conserving existing forests in developing countries. The methodology also provides a means of selecting the most suitable scale at which models of carbon risk can be parameterised through a comparison of the results of models parameterised with data from different spatial scales. In this study three regions were defined by climatic, demographic and land use characteristics (Highlands, Cañadas and Selva). Our results indicate that this division of the study area was appropriate for two of the three regions. However, the Selva region required further sub-division to take account of the marked variations in the land tenure and underlying processes affecting land-use change in this region. While it is possible that similar improvements could be made by sub-dividing the other regions the errors are already relatively low: less than 5% of vulnerable carbon stocks when applied to areas in excess of 10,000ha and around 10% at a scale of 2,500ha. In the example of the La Corona community with an area of 2,274ha a risk buffer of 11.5% of vulnerable carbon stock was set based on the performance of the DistRd model at the 2,500ha scale. If the project were expanded to cover 10,000ha or more then this risk buffer could be reduced to 1% of vulnerable carbon stocks.

Acknowledgements

This publication is an output from a research project funded by the United Kingdom Department for International Development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of DFID. Project R7274 Forestry Research Programme.

References

Castillo, M.A., Hellier, A., Tipper, R. and de Jong, B.H.J. 2002. Carbon emissions from land-use change: a regional analysis of causal factors in Chiapas, Mexico. Submitted

De Jong, B.H.J., Ochoa-Gaona, S., Castillo-Santiago, M.A., Ramírez-Marcial, N. and Cairns, M. 2000. Land-use/land-cover change and carbon flux in the Selva Lacandona, Mexico. *Ambio* 29 (8): 504-511.

Stuart, M.D. and MouraCosta, P. 1998. *Climate Change mitigation by forestry: a review of international initiatives.* Policy that works for forests and people series no.8. International; Institutes for Environment and Development, London.

UNFCC 2000. Uniform Reporting Format: Activities Implemented Jointly Under the Pilot Phase Costa Rica - Norway Pursuant to Documents FCCC/CP/1995/7/Add.1 Decision 5/CP.1, para. 2 (b) June 2000 http://unfccc.int/program/coop/aij/aijact01/crinor01-00.html

Appendix 1: List of carbon risk matrices used in the analysis for DistRd and DistAg models. Figures in cell are the expected deforestation over the next 21 years as a percentage of current forest area.

Study area

		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	52.1%	43.2%	27.5%
	>15 to 30	43.6%	36.6%	28.0%
	>0 to 15	36.8%	27.9%	20.9%
	0	32.1%	19.7%	5.8%

Highlands

		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	44.9%	34.8%	17.3%
	>15 to 30	38.5%	35.1%	26.0%
	>0 to 15	32.9%	29.9%	18.0%
	0	30.4%	24.3%	15.7%

Cañadas

		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	57.4%	44.2%	38.7%
	>15 to 30	45.9%	35.0%	31.1%
	>0 to 15	38.1%	26.1%	19.6%
	0	30.4%	21.4%	10.2%

Selva

		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	63.6%	51.4%	24.8%
	>15 to 30	47.9%	41.8%	25.8%
	>0 to 15	39.8%	29.7%	22.0%
	0	33.0%	18.7%	5.2%

Marques

		Distance from roads (m)		
		0-1000	1000-2000	>2000
Population density (hab/km2)	>30	65.2%	41.8%	18.9%
	>15 to 30	48.1%	32.1%	24.8%
	>0 to 15	42.0%	27.5%	23.3%
	0	36.0%	21.8%	15.6%

Study area		**Distance from agriculture (m)**		
		0-500	500-1000	>1000
Population density (hab/km2)	>30	55.4%	45.6%	48.4%
	>15 to 30	53.0%	42.1%	37.6%
	>0 to 15	51.6%	38.0%	26.9%
	0	49.2%	35.9%	11.4%

Highlands		**Distance from agriculture (m)**		
		0-500	500-1000	>1000
Population density (hab/km2)	>30	51.3%	34.9%	35.6%
	>15 to 30	48.9%	33.2%	33.5%
	>0 to 15	52.3%	37.7%	24.2%
	0	52.5%	43.0%	20.3%

Cañadas		**Distance from agriculture (m)**		
		0-500	500-1000	>1000
Population density (hab/km2)	>30	61.1%	54.6%	53.3%
	>15 to 30	53.3%	43.4%	40.2%
	>0 to 15	49.9%	35.6%	27.4%
	0	43.1%	32.4%	15.4%

Selva		**Distance from agriculture (m)**		
		0-500	500-1000	>1000
Population density (hab/km2)	>30	74.9%	61.4%	46.3%
	>15 to 30	65.8%	57.3%	36.0%
	>0 to 15	58.7%	46.0%	27.6%
	0	55.6%	37.9%	10.4%

Marques		**Distance from agriculture (m)**		
		0-500	500-1000	>1000
Population density (hab/km2)	>30	91.0%	77.2%	40.6%
	>15 to 30	99.4%	93.2%	35.5%
	>0 to 15	66.2%	75.0%	31.3%
	0	43.9%	25.4%	24.4%