INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0800
THE USE OF REMOTE SENSING AND GEOGRAPHIC INFORMATION SYSTEMS FOR SOIL EROSION HAZARD MAPPING IN CHIAPAS, MEXICO

by

Robert E. Currie

A thesis submitted in conformity with the requirements for the degree of Master of Sciences
Graduate Department of Geography
University of Toronto

© Copyright by Robert Emmet Currie 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-45846-6
Robert E. Currie
The Use of Remote Sensing and Geographic Information Systems for Soil Erosion Hazard Mapping in Chiapas, Mexico
Master of Sciences, 1997
Department of Geography
University of Toronto

Abstract

Land degradation due to soil erosion is a serious problem in the central highlands of Chiapas, Mexico, and one that is of interest to environment-conflict researchers. It is one of many environmental scarcities which may have contributed to violent insurgency in the area in 1994. A soil erosion model, SEMMED, which incorporates satellite images and geographic information systems (GIS) is used to model erosion hazard at a regional scale. Satellite images are used to derive vegetative inputs to the model for the 900 km² study area. Soils, precipitation and other model inputs are entered into the GIS. Erosion predictions are made by calculating soil loss due to splash detachment and transport capacity and the lesser of the two values taken as the limiting factor in the erosion process. Data inputs and model limitations are discussed. Finally, the utility of remote sensing and geographic information systems to environment-conflict research is reviewed.
Acknowledgements

For their generous assistance in the preparation of this thesis I would like to thankfully acknowledge the following people: Larry Band, Ferko Csillag, Tad Homer-Dixon, Thomas de Jong, Phillip Howard, Karen O'Brien, Dr. Evon Zogt, Dr. George Collier, Trena Creed, Tim Ziegler and Carlos Rodriguez.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Statement of Problem</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Research Questions</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Discussion of Research Questions</td>
<td>9</td>
</tr>
<tr>
<td>1.5 Approach to Thesis</td>
<td>11</td>
</tr>
<tr>
<td>2.0 Environmental Conditions in the Study Area</td>
<td>12</td>
</tr>
<tr>
<td>2.1 Study Area Selection</td>
<td>12</td>
</tr>
<tr>
<td>2.2 Vegetation and Land Use</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Natural Vegetation</td>
<td>14</td>
</tr>
<tr>
<td>2.2.2 Agriculture</td>
<td>15</td>
</tr>
<tr>
<td>2.2.3 Land Use Patterns</td>
<td>17</td>
</tr>
<tr>
<td>2.3 Physical Characteristics of the Study Area</td>
<td>19</td>
</tr>
<tr>
<td>2.3.1 Topography</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2 Climate and Precipitation</td>
<td>20</td>
</tr>
<tr>
<td>2.3.3 Soils</td>
<td>21</td>
</tr>
<tr>
<td>3.0 Spatial Data Models</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Morgan, Morgan and Finney Soil Erosion Model</td>
<td>23</td>
</tr>
<tr>
<td>3.2 SEMMED</td>
<td>23</td>
</tr>
<tr>
<td>4.0 Data Sets</td>
<td>25</td>
</tr>
<tr>
<td>4.1 Satellite Images</td>
<td>25</td>
</tr>
<tr>
<td>4.2 INEGI Maps</td>
<td>28</td>
</tr>
<tr>
<td>4.3 Precipitation Data</td>
<td>29</td>
</tr>
<tr>
<td>4.4 Other Data</td>
<td>30</td>
</tr>
<tr>
<td>5.0 Deriving the Vegetative Model Inputs</td>
<td>32</td>
</tr>
<tr>
<td>5.1 Supervised Classification</td>
<td>32</td>
</tr>
<tr>
<td>5.2 Unsupervised Classification</td>
<td>34</td>
</tr>
<tr>
<td>5.3 NDVI Segmentation</td>
<td>35</td>
</tr>
<tr>
<td>5.4 Further Processing and Observations</td>
<td>37</td>
</tr>
<tr>
<td>6.0 Spatial Data Processing</td>
<td>44</td>
</tr>
<tr>
<td>6.1 Building the Spatial Database</td>
<td>44</td>
</tr>
<tr>
<td>6.2 Determining Splash and Transport Capacity</td>
<td>46</td>
</tr>
<tr>
<td>6.3 Calibrating the Overland Flow Values</td>
<td>53</td>
</tr>
<tr>
<td>7.0 Results and Observations</td>
<td>56</td>
</tr>
<tr>
<td>7.1 Erosion Hazard Maps</td>
<td>56</td>
</tr>
<tr>
<td>7.2 Model Sensitivities</td>
<td>64</td>
</tr>
<tr>
<td>7.3 Model and Methodology Shortcomings</td>
<td>69</td>
</tr>
<tr>
<td>8.0 Conclusions</td>
<td>73</td>
</tr>
<tr>
<td>8.1 Discussion of Research Questions</td>
<td>73</td>
</tr>
<tr>
<td>8.2 Conclusion</td>
<td>77</td>
</tr>
</tbody>
</table>
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Landsat MSS data description</td>
<td>26</td>
</tr>
<tr>
<td>Table 2</td>
<td>Vegetation model input calculation</td>
<td>41</td>
</tr>
<tr>
<td>Table 3</td>
<td>Input sensitivity testing</td>
<td>65</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Schematic of approach to thesis</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2</td>
<td>The study area in Chiapas, Mexico</td>
<td>13</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Land classification by altitude</td>
<td>14</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Milpa plots</td>
<td>16</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Severely eroded slopes</td>
<td>18</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Digital elevation model</td>
<td>19</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Monthly precipitation levels</td>
<td>21</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Major soil types in the study area</td>
<td>22</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Composite MSS image of study area</td>
<td>25</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Normalized difference vegetation index</td>
<td>27</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Soils map of region</td>
<td>28</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Rainfall as a three dimensional surface</td>
<td>30</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Cluster merging</td>
<td>35</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Land cover classes</td>
<td>36</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Land cover class breakdown</td>
<td>38</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Maize cultivation and pasture</td>
<td>40</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Landscape of study area</td>
<td>42</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Flow chart of splash detachment calculation</td>
<td>46</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Results of splash detachment calculation</td>
<td>48</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Flow chart of transport capacity calculation</td>
<td>49</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Reclassified rooting depth values</td>
<td>51</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Results of transport capacity calculations</td>
<td>53</td>
</tr>
<tr>
<td>Figure 23</td>
<td>Erosion predictions for 1980 image</td>
<td>56</td>
</tr>
<tr>
<td>Figure 24</td>
<td>Erosion predictions for 1980 image</td>
<td>57</td>
</tr>
<tr>
<td>Figure 25</td>
<td>Erosion predictions for 1992 image</td>
<td>57</td>
</tr>
<tr>
<td>Figure 26</td>
<td>Erosion predictions for 1992 image</td>
<td>58</td>
</tr>
<tr>
<td>Figure 27</td>
<td>Predictions using saturated infiltration capacity</td>
<td>58</td>
</tr>
<tr>
<td>Figure 28</td>
<td>Distribution of results by classification method</td>
<td>59</td>
</tr>
<tr>
<td>Figure 29</td>
<td>Difference image between classification method</td>
<td>60</td>
</tr>
<tr>
<td>Figure 30</td>
<td>Generalized representation of results</td>
<td>62</td>
</tr>
<tr>
<td>Figure 31</td>
<td>Splash detachment limited results</td>
<td>63</td>
</tr>
<tr>
<td>Figure 32</td>
<td>Sensitivity testing methodology</td>
<td>66</td>
</tr>
<tr>
<td>Figure 33</td>
<td>Lower rooting depth impact on results</td>
<td>67</td>
</tr>
<tr>
<td>Figure 34</td>
<td>Detachability index impact on results</td>
<td>68</td>
</tr>
</tbody>
</table>
1.0 Introduction

1.1 Background

The notion that resource scarcity can lead to acute conflict is neither counter-intuitive nor new. Understanding more precisely the nature of the relationship, however, has become an issue of increased interest as environmental problems become more profound in areas of the world that are prone to violent outbreaks. Environment-conflict researchers believe the contribution of environmental scarcity to certain instances of civil or international strife are more significant than previously believed.

Environmental scarcity in this context is broadly grouped into three categories: supply-induced, demand-induced and structural (Homer-Dixon et al., 1993). Supply-induced environmental scarcity refers to the depletion of physical natural resources. This can refer to non-renewable resources such as oil and minerals, however, it is also commonly caused by over-exploitation of renewable resources such as forests or ground-water.

Demand-induced environmental scarcity is caused by an increasing human population sharing a finite resource base. As the denominator grows (population), the per capita share of a finite resource shrinks. The third type of environmental scarcity identified is structural. Disproportionate access to resources by elite, powerful groups in a society represent this type of scarcity (Homer-Dixon et al., 1993).

These scarcities can contribute to violence in a variety of ways. The Project on Environmental Change and Conflict outlines three principle ways of categorizing such
conflicts: 1) ‘simple scarcity’ conflicts which arise directly due to depletion or degradation of one or many resources, 2) environmental stress-induced population migrations leading to ‘group identity clashes’, and, 3) ‘deprivation conflicts’ resulting from the economic decline and weakened social institutions, in turn caused by resource depletion and scarcity (Homer-Dixon, 1994).

For a host of reasons, these types of environmental scarcity are more likely to occur in developing countries (Hauchler, 1994, Homer-Dixon, 1991). Clearly, structural environmental scarcity will occur more frequently in non-democratic societies where elite groups gain access to resources at the expense of less powerful ones. This tends to occur particularly when there exists some type of cultural, tribal or religious cleavage between those in power and other minority groups. Since most of the world’s dictatorships and non-democratic governments exist in the developing world, this is where structural environmental scarcity is most frequently seen.

Demand-induced scarcity, as a function of population growth, is a serious and worsening problem in many less developed countries. From 1995 to 2025, it is projected that the population increase in North America (excluding Mexico) and Europe combined will be less than 100 million. By contrast, the populations of Africa, Asia and Latin America are expected to increase by over 2.5 billion over the same time period (World Resources, 1994. pp. 268-269). For the latter areas, this constitutes a population increase of 55%
and a corresponding increase in competition for an already scarce, and often shrinking, resource base.

In many less developed countries, the greatest challenge will be feeding people and providing basic sustenance. Most of the world's arable land that can be put into agricultural production has been, and the yield per unit area increases that have been witnessed due to fertilizers and pesticides cannot continue at the same rate (WRI, 1991 p. 84). With limited potential to put more land into agricultural production, and limited increases in yields per unit area, increased population translates to a declining amount of arable land per capita (WRI, 1991, 1994, Homer-Dixon, 1993). This will lead to increased competition for this life-sustaining resource, and forced migration for groups which are denied an adequate supply.

Supply-induced environmental scarcities are not necessarily more common in the developing world. However, their ramifications tend to be more profound in developing economies than in developed ones, most of which have long since transitioned into tertiary and quaternary industry. The immediate well-being or even survival of local populations is much more tightly linked to the resource base in the developing nations where entire communities may rely on the exploitation of a single resource. Access to water, for example, has long been the source of conflict in many regions where irrigation farmers, industry and individuals compete with upstream users, diverters and polluters (Gleick, 1993, pp. 97-99).
The research does not contend, however, that environmental scarcity is either sufficient to cause violent conflict or is necessary for it to occur. Rather, it is through the study of violent conflicts where environmental degradation is clearly evident that potential causal relationships between violence and environmental scarcity can be assessed. A complex of factors combine to create a climate for violence. They may include cultural norms for conflict resolution, the extent of a group's enfranchisement in the political system, levels of economic development and well-being, ethnic or religious rivalries, access to health care and the perceived legitimacy of the state.

If acute environmental scarcities contribute to violent conflict it will likely be in conjunction with a set of other political, social and economic factors which jointly create the climate for violence. They may, as in the case of a simple scarcity conflict, be identifiable as a contributing factor to violence in their own right. They may also be seen as more distant contributors to these other inputs to a volatile climate. For example, a lack of arable land due to population increase and/or land degradation will contribute to economic decline among peasant farmers, which may in turn enhance the solidarity among this group in society, strengthen their sense of disenfranchisement and heighten their grievance towards government.

In attempting to assess the relationship between various types of environmental scarcity and acute conflict, current and historic data sets are required which either quantify or assist in quantifying the resource base. The monitoring of environmental indicators and
resources, and the ability to discern the rate of degradation or exploitation of resources, is of great importance to the environment-conflict researcher. However, current, reliable, regional-scale data sets for natural resources often don’t exist in areas of significance. Much of this research is currently directed towards the developing world where the availability of accurate data sets is often found wanting.

For example, in the Chiapan case study, a number of sources indicated that soil erosion in the highlands is a serious problem, however no hard data sets existed to corroborate that claim or quantify it. The FAO/UNESCO world soil erosion map series from 1990 is at a scale of 1:10 000 000 and, therefore, gives only highly aggregated erosion indicators using broad categories.

1.2 Statement of Problem

The state of Chiapas, Mexico has become known to many as a “rich land and a poor people” (Benjamin, 1996). A relative abundance of natural resource from forests to agricultural land to water for hydroelectric power generation is not evident in the standard of living of the average citizen. In addition to being one of Mexico’s richest states in natural resources, it is also the state with the highest proportion of indigenous peoples, with 22% (INEGI, Census 1994). While the numbers of indigenous are significant, they are not politically enfranchised. They typically have little if any formal education, and engage in subsistence agriculture in sparsely populated regions of the
The ruling Partido Revolucionario Institucional (PRI) in Mexico has maintained power since the revolution in 1910 and has had success in Chiapas despite its grossly unequal treatment of the state's Mayan descendants.

On New Year's Day, 1994, the Ejercito Zapatista Liberacion Nacional (EZLN, or Zapatista National Liberation Army) emerged from the forests and jungle to take control of six cities, including the state's second largest, San Cristobal de las Casas. Declaring war on the Mexican government, the EZLN demanded, among other things, land, education and justice for the indigenous peoples of Chiapas and Mexico in general. The Mexican government and army were caught by surprise, and although they quelled the insurrection within a few days, it was not before the entire world knew about conditions in Chiapas and the struggle of the EZLN. The EZLN withdrew into the forests without surrendering and have been engaging in peace talks with the government for the past three years, though there have been few concrete results.

This is an interesting case study for environment-conflict researchers because all three types of environmental scarcity are clearly evident. Structural scarcity, insofar as the indigenous were concerned, is evident and it is well documented. Also well documented are the extremely high fertility rates among indigenous populations, and population growth further augmented in the 1980's by the migration of Guatemalan refugees fleeing persecution through the Peten jungle into Chiapas (Collier, 1994, Benjamin, 1996, Howard et al., 1996). Numerous sources indicate that, as far back as the early 1970s,
agricultural practices among the Maya in the central highlands were unsustainable and that more and more forest was being cleared to compensate for eroded lands which could no longer sustain crop growth (Cisneros, 1986, Zogt, 1969, Collier, 1975, 1995). Data which quantified the loss of the forest resource, for example, were not available.\(^5\) Additionally, the extent to which soil erosion had negatively impacted local farmers was not available. In assessing grievances and the perception of opportunity through insurrection, the ability to estimate economic hardship that is directly related to resource depletion and/or degradation is critical.\(^6\) Furthermore, in this case, the insurgents were very clear in specifying a lack of access to agricultural land as a major grievance, thus heightening interest in land-related data sets.

Data which quantified the loss of the forest resource, for example, were not available. Additionally, the extent to which soil erosion had negatively impacted local farmers was not available. This information is of great importance because of the multicausal nature of violent conflict. Increases in soil erosion and land degradation lead directly to declines in agricultural productivity and output.\(^7\) This, in turn, creates economic hardship and may contribute to the sense of grievance towards the state on the part of the peasant farmers who are adversely effected by the resource degradation. Assessing the contribution that environmental factors make to grievance and how they interact with other causal agents requires supporting data sets. There is a great demand for data sets that are suitable for such analysis and yet researchers find the supply inadequate.
The primary focus of this thesis is the creation of soil erosion hazard maps for the study area in the central highlands of Chiapas. This will be done by adapting a soil erosion model (SEMMED) that incorporates the use of satellite imagery and a geographic information systems (GIS). The results will be maps indicating the locations in the study area which are prone to severe land degradation through soil erosion.

1.3 Research Questions

In order to assess soil erosion in the Chiapan study area, the utility of remote sensing (RS) and GIS technology in regional-scale erosion hazard mapping will be investigated and the following research questions addressed:

1. What are the spatial-resolution imposed limitations of using Landsat Multi Spectral Scanner (MSS) data for the type of vegetation analysis required by this and other soil erosion models?

2. What are temporal-resolution imposed limitations of using RS data for studying land-cover change in a two-season climate such as Chiapas’?

3. Can MSS data be useful for assessing vegetative cover (in the way required by this study) in the absence of field data to assist in image analysis and substantiate findings?

4. Is the SEMMED model easily adaptable to the Chiapan climate, given that it is designed for a Mediterranean environment? What modifications will be required?

5. Can Remote Sensing and GIS be useful technologies in environment-conflict research?
1.4 Discussion of Research Questions

1. Landsat MSS data is some of the cheapest and most readily available satellite imagery that is available today. For change detection studies, it is also the data set that dates back the furthest, which is an advantage. However, the 79mx79m spatial resolution has limitations and is not necessarily adequate for assessing land cover in landscapes with high spatial variability. For this and other soil erosion models, land cover classes will need to be derived from the remotely sensed data and the suitability of MSS data for this purpose will be assessed.

2. Vegetation is differentiated from other features in satellite images based on its spectral signature, or the way it reflects different parts of the electromagnetic spectrum. The signature for healthy vegetation is heavily influenced by its chlorophyll absorption (Mather, 1987, Lo, 1986). Seasonal vegetation such as crops or deciduous forests will appear entirely differently from one season to the next where there is high variability in foliage, among other things. For analyzing land cover change over time, capture dates of satellite images will determine the extent to which they can be useful. Time-series satellite images with different capture dates will be used in this study and this question addressed.

3. One way to ‘classify’, or define land cover classes, in a satellite image is to refer image pixels at specific locations to field data collected at the same spot. This process
assists in ‘training’ the image processing software to recognize patterns in the combination of bands that are being analyzed. Without a priori ground knowledge, determination of land cover classes must be done by other means. The viability of these means will be assessed.

4. While this modeling exercise follows the same process that Simon de Jong (1994) outlined for adapting the Morgan, Morgan and Finney soil erosion model (1984) to work on a distributed basis, there are several differences in this study. Data inputs vary, for example, all vegetative inputs are derived here by different means. Hydrological routing algorithms differ, although perform similar tasks. Additionally, the model requires a number of inputs that are climate-specific and these will have to be tuned to work with the Chiapan climate.

5. An important source of data that is increasingly accessible to researchers are remotely sensed images from the Landsat and SPOT series of satellites. Originally launched in 1972 and 1986 respectively, they have been gathering and archiving massive amounts of data covering most of the earth’s surface at different spatial and temporal resolutions for over two decades. Change detection studies and resource monitoring over time is possible using these data sets and commonly available image processing software and geographic information systems. The potential advantages and pitfalls of these data for environment-conflict research will be addressed.
1.5 **Approach to thesis**

The following diagram outlines the approach and sequence of the various stages of this thesis.

![Diagram of approach to thesis]

- Remote sensing techniques and theory
- Soil erosion physical processes
- Environment-conflict research and theory

**Data Acquisition**

**Building Spatial Database**

**Spatial Data Processing**
- Image classification
- Erosion modeling

**Modifications to SEMMED**
- Production of erosion hazard maps

**Evaluate pertinence of study to**
- Environment-conflict research

**Discussions and Conclusions**

*Figure 1. Schematic of approach to thesis.*
2.0 Environmental Conditions in the Study Area

2.1 Study Area Selection

In a broad sense, the study area was to be the central highlands of Chiapas, Mexico. This geographic area is pertinent because it is the power base for EZLN and the location of their uprising in 1994. A specific, and more geographically limited, area to study had to meet the following criteria:

1. The study area must be in the central highlands of Chiapas, and preferably in an area known to have EZLN support and activity.
2. The Landsat images that were available must address the area with minimal, if any, cloud cover in any of the three time-series satellite images.
3. The area should be geographically coincident with the available hard-copy maps for the region.
4. The area should contain some form of human settlement and other human activity such as agriculture, as these areas are likely to exhibit soil erosion and land cover change over time.

The study area selected is approximately 30km by 30km and has the town of San Cristobal de las Casas, in the south-west corner. This area meets all above requirements and is an area likely to have experienced significant change over the time period covered by the satellite images, as San Cristobal is the second largest city in Chiapas and the most important in the region. Furthermore, the study area was part of the area under the control of the EZLN during the 1994 insurrection. Additionally, I was based in San Cristobal for several weeks in 1995. As an international observer at the second round of peace talks between the EZLN and the Mexican government, I spent time in San
Lorenzar, the site of the talks, and also in Chamula and Zinacantan, the focus of most of the work done by Zogt (1969, 1994) and Collier (1975).

Figure 2. The study area in Chiapas, Mexico.

The seminal research on the peoples of the Chiapan highlands was carried out by various people through the Harvard Chiapas Project which began in 1964. A good deal of the research centered on the municipios and centres of Zinacantan and Chamula, both of which are within a few kilometres of the study area for this project. These communities are located on the same plateau as this project's study area, and share culture, language, social organization as well as agricultural practices with communities. These similarities have been extremely valuable in interpreting various data sources.
2.2 Vegetation and Land Use

Figure 3. Red is ‘Tierra Fria’, yellow is ‘Tierra Templada’, and green is ‘Tierra Caliente’. Cyan is no data value.

2.2.1 Natural Vegetation

As precipitation and temperature patterns are linked to elevation, so is the vegetation found in the study area. Zogt describes the Mayan terminology for land classes as being Cold Country (Tierra Fria), below 900m, Temperate Country (Tierra Templada), from 900 to 1800m, and Hot Country (Tierra Caliente) which is over 1800m (Zogt, 1969). Figure 3 shows how the study area is segmented according to these categories. Cold Country hosts broad-leafed, non-deciduous pine and oak forests, although generally only on slopes which render agriculture impossible. “Mosses and epiphytes, such as bromeliads, cling to trees there, preserved by the misty cloud cover that tends to shroud the higher peaks” (Collier, p.25). In Temperate Country, mixed forests are present and
contain both coniferous and deciduous species, however the canopy remains dense as does the understory. In the lower elevations towards the east of the study area climax forest will be tropical rainforest, although any climax stage tropical rainforest that still exists is likely to be in the Montes Azul Biosphere Reserve, further east into the Selva Lacondon.

Other than forest, land cover will be either crops, grassland or scrub growth. Scrub growth in various stages is evident as fallow fields revert back towards forest. Depending on the fertility of the land and the altitude, follow field may be taken over by grasses for a period of time before shrubs begin to grow again. Regeneration in the cooler highlands will take longer and the stages of grass, scrub and ultimately forest taking over fields left in fallow are more pronounced there than in the lowlands (Collier, 1975, p25). Throughout the study area, land which is no longer fertile will revert to grassland and eventually will be unable to support vegetation at all.

2.2.2 Agriculture

Agriculture in this area is dominated by milpa production, which refers to the cultivating of maize and beans together in the same fields. The seeds are planted in the same holes and the faster growing corn stalk provides support for the bean plant’s growth. Again altitude is instrumental in dictating the timing of agricultural activities. The maize growing season is typically from May to October, corresponding with the rainy season. In the highlands around San Cristobal the maize is harvested by the end of October, but
harvesting may continue into December and January at the lower, more temperate elevations (Vogt, 1969). Agriculture is practiced on small, privately owned plots of land and larger ‘ejido’s’ which are communally owned farms or ranches. Larger-scale agriculture also takes place on land that is generally in the lowlands and is privately held by ‘latifundistas’ or ‘rancheros’ (Howard, 1996, Collier, 1995, Vogt, 1969).

Figure 4. Milpa plots are seen here around a small settlement (From Zogt, Zinacantan, 1969).

Population-induced land scarcity and declining land fertility in the highlands make it difficult or impossible for most highland farmers to provide for their families without lowland farming either in the eastern lowlands or to the west in the Grijalva River valley. The later growing season in the lowlands allows the Tzotzil farmers to work rented lowland plots in addition to their own. This works on a tithe system which is based on the acreage being rented. The Tzotzil or Tzeltal farmers guarantee a certain amount of
corn as rent for using the land, and thus appropriate the risks in the case of a poor harvest.

While maize is definitely the primary crop grown in the area, beans are an important second for their provision of protein to the Tzotzil and Tzeltal diet which only rarely includes meat. Around human settlements in the highlands, more intensive agriculture is practiced and fertilizers and pesticides are used to alleviate the need to let land lie fallow. In these areas, cash crops like squash, chili and chayote are grown. In the lowlands, coffee and bananas are cultivated in addition to milpa (Benjamin, 1996, Collier, 1975).

2.2.3 Land Use Patterns

In *Fields of the Tzotzil*, Collier (1975) describes the stages of succession on the Chamula landscape as they range from climax forest to rocky eroded land that cannot support vegetation at all. Chamula lies about 10 kilometres to the north-west of the study area but for reasons of altitude and similar population-induced pressure on the land, provides useful insight into what has happened elsewhere in the highlands. Climax forest, the first stage, occurs infrequently and could therefore be expanded to include second generation forest growth on land which has not been farmed for generations. The second stage takes place when the forest is cleared, burned and cultivated for a year or two before being allowed to revert to forest. This stage can be seen from the air as a small plot of cleared land surrounded by forest and is more likely to occur as highland farmers select tracts of land to farm away from their settlements (Vogt, 1969). In the third stage, more
intensified swidden agriculture takes place over larger, contiguous tracts of land. Typically this will be surrounding or nearby to a human settlement and delineation of land parcels is evident either by cultural or physiographic features. Sheep herding is introduced and controlled systematically such that fields are fertilized by sheep and do not require periods of fallow.

The intensity of agriculture in the third stage is not sustainable and decreases the fertility of the land. Increasing amounts of acreage is useful only for pasture in the fourth stage and sheep herding predominates. Settlement farmers must travel greater distances to tend their crops and eventually the settlement will move on. The fifth and sixth stages trace the gradual decline from grassland eroded around trail systems and steep slopes to rocky, bare soil that sustains no growth at all.

Figure 5. Severely eroded slopes in the Chiapan highlands. (From Collier, Fields of the Tzotzil, 1975)
2.3 Physical Characteristics of Study Area

2.3.1 Topography

The study area is located in the central highlands of Chiapas, slightly north-west of the state’s geographic centre. It lies between 16 45 00 and 17 00 00 degrees north of latitude and between 92 30 00 and 92 45 00 degrees west of longitude. The relief is substantial, as elevation ranges from 750m to 2900m above sea level and the topography is extremely rugged. San Cristobal and its surroundings in the south-west corner of the study area are some of the highest points in the state and elevations drop off dramatically towards the north-east as the landscape trends from the central highlands to the eastern lowlands.

The central plateau, on which San Cristobal is found, is a limestone uplift broken through by volcanic extrusions (Collier, 1975, p.21). The terrain is characterized by small to medium sized valleys with often steep slopes giving way to flat valley bottoms.

Figure 6. The digital elevation model of the study area. Elevation is in metres.
2.3.2 Climate and Precipitation

The Atlas Nacional de Mexico describes the area as an intertropical convergence zone and breaks the study area into two main climatic areas: tropical savanna in the south-west third and tropical monsoonal for the north-east two thirds. These climates are characterized by two distinct seasons, a wet and a dry one that each last approximately six months. The wet season runs from May to October and the dry season from November to April. Mean annual precipitation levels in the study area range from 1400mm to 1800mm and shows a generally inverse relationship with altitude. This is primarily due to the convective rains due to higher temperatures in the lowlands. Conversely, temperature is quite consistent throughout the year, with average monthly means ranging from about 16°C in December and January and 20°C for the months of June through September. Elevation is influential in temperature patterns and the lowlands have higher temperatures and more rainfall than the highlands. In the lower elevations, precipitation exceeds evapotranspiration 11 or 12 months of the year, whereas the number is 9 or 10 months at the higher elevations around San Cristobal (Atlas Nacional de Mexico, 1996). Collier notes that during the rainy season, patterns of rainfall occur locally to the many valley systems in the area, with afternoon warming creating convection patterns culminating in late-afternoon storm bursts.
Figure 7. Monthly precipitation averages in the Chiapan highlands.

2.3.3 Soils

The INEGI soils map acquired for this project uses the FAO/UNESCO soils classification system. Soils in the study area are influenced by the limestone volcanic extrusions which form the central highlands and are generally fertile, although prone to erosion (Vogt, 1969, p.4). Soils in the study area are as follows: (INEGI soils map 1:250000, 1986, Atlas Nacional de Mexico 1: 8000000, 1996)

1) Chromic Luvisols. Luvisols generally have an alluvial accumulation of clay. They display an argillic B horizon and have base saturation of ≥ 50%.

2) Gleysols. These soils are formed from unconsolidated materials which have hydromorphic properties. In the study area, the gleysols are found on the San Cristobal plateau, which is a characteristic location for this soil type. They exhibit poor drainage and often show cambic B horizons.
3) Rendzina. A mollic A horizon rich in organic matter generally overlying a calcareous material, which in this case is limestone. They are no more than 50 cm thick.

4) Haplic phaeozem. These phaeozems are associated with the forest steppes and have a high organic content and a deep leaching of calcium carbonate. They tend to display a mollic A horizon and argillic B horizon with a base saturation of $\geq 50\%$ and no calcic or gypsic horizon.

5) Humic acrisols. Base saturation is $\leq 50\%$ and has an argillic B horizon.

![Figure 8. Major soil types in the study area.](image-url)


3.0 Spatial Data Models

3.1 Morgan, Morgan and Finney Soil Erosion Model

The Morgan, Morgan and Finney (MMF) model was developed with field-size areas on hillslopes in mind. It attempts to preserve the simplicity of the Universal Soil Loss Equation (USLE) while utilizing more recent approaches and advances in understanding the soil erosion process (Morgan, Morgan, Finney, 1984). The model breaks down soil erosion into two phases: the water phase and the sediment phase. The water phase determines kinetic energy of rainfall based on intensity and volume of precipitation and the transport capacity of soil by overland flow. The sediment phase is concerned with the detachment rate of soil (from kinetic energy determined in the water phase) and the transport capacity of overland flow. The model compares the values obtained for splash detachment and transport capacity and accepts the limiting factor, or lesser of the two. According to Morgan, Morgan and Finney (1984), the model is most sensitive to rainfall and soil parameters when erosion is transport-limited and to rainfall and precipitation interception when erosion is detachment-limited. This makes soils and rainfall data particularly important to accurate forecasting (Morgan, 1986).

3.2 SEMMED

SEMMED, or the Soil Erosion Model for Mediterranean Climates, was developed by Steven de Jong and described in his 1994 work, Applications of Reflective Remote Sensing for Land Degradation Studies in a Mediterranean Environment. In this de Jong details the development of the model for his study area in Bas-Vivarais of Ardeche
Province, France. SEMMED has as its physical basis the MMF soil erosion model. As mentioned, the MMF model was designed to assess soil erosion on field-size plots and does not consider spatially differentiated data. The use of RS images and a GIS make SEMMED a distributed model whereby the spatial distribution of input variables is incorporated into the processes being modeled.

Image processing is used in SEMMED to correlate field data and spectral indices from Landsat TM images in order to obtain values for vegetation-related inputs to the model. The precipitation interception factor (P), crop cover factor (C) and the ratio of actual to potential evapotranspiration (Et/Eo) are obtained using such techniques. De Jong collected field data at 33 test plots and statistically correlated his field data with a number of spectral indices for vegetation. Using Landsat Thematic Mapper (TM) images, this methodology allowed reasonably accurate vegetation inputs to be derived for a much larger study area than was covered by his field work. SEMMED also uses a DEM and a slope map derived from it to create a channel map which indicates the directional tendencies of water to flow and accumulate. This is the basis, along with the DEM, for creating a distributed flow map of annual runoff where it exceeds the soil's ability to absorb.
4.0 Data sets

4.1 Satellite Images

The satellite images were obtained through the North American Landscape Characterization (NALC) program. This project is part of the National Aeronautics and Space Administration (NASA) Landsat Pathfinder Program and has rendered public domain Landsat Multispectral Scanner (MSS) images for most of the North American continent. The NALC triplicate data sets were assembled for land cover change analysis and include MSS images for a given scene from the 1970s, 1980s and 1990s (NALC Research Brief, 1995). MSS scenes are indexed by a row and path number, and the study area for this thesis is taken from the scene identified as path 48, row 21. Images were received for the following dates for this scene: February 15, 1974, November 17, 1980, November 25, 1984 and August 27, 1992. Two images were supplied for the 1980s to ensure that, between the two images, all areas in the scene are cloud-free.

Figure 9. A composite MSS image of the study area using bands 4,2,1.
Remotely sensed images are a matrix of digital numbers representing reflectance values in a certain band or range of the electromagnetic spectrum. An image, such as the Landsat images used in this project, contain several layers of information, each one representing the reflectance values in a particular band of wavelength. Landsat MSS data comes in four bands which cover the visible and near-infrared portion of the electromagnetic spectrum. Table 1 summarizes MSS data.

<table>
<thead>
<tr>
<th>Band</th>
<th>Wavelength</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1</td>
<td>0.50 - 0.60 μm</td>
<td>Green: Scans the region between the blue and red chlorophyll absorption bands.</td>
</tr>
<tr>
<td>Band 2</td>
<td>0.60 - 0.70 μm</td>
<td>Red: Chlorophyll absorption band of healthy green vegetation.</td>
</tr>
<tr>
<td>Band 3</td>
<td>0.70 - 0.80 μm</td>
<td>Reflective infrared: Responsive to the amount of vegetation biomass.</td>
</tr>
<tr>
<td>Band 4</td>
<td>0.80 - 1.10 μm</td>
<td>Reflective infrared: Useful for vegetation surveys and penetrating haze.</td>
</tr>
</tbody>
</table>

Table 1. Landsat MSS data description. (Mather, ERDAS, Harris)

For each scene, and for each decade, the data sets include the four MSS bands, a Normalized Differential Vegetation Index (NDVI) image, a pixel identification image (in the case that the image provided is a composite image made to minimize cloud cover) and a digital elevation model (DEM). While the spatial resolution of MSS data at capture is 59m X 79m, these images have been reprojected and resampled to a 60m x 60m pixels in a UTM projection (NALC Image Processing, 1995).
The use of vegetation discrimination in remotely sensed images is assisted by the fact that vegetation has low reflectance in the visible spectrum and high reflectance in the near-infrared range of the electromagnetic spectrum (Tucker, 1979). In analyzing remotely sensed images, there are several commonly used indices that are used for specific purposes. For example, the tasseled cap transformation and the NDVI are two such indices used commonly in vegetation analysis. The NDVI is a ratio transform which accentuates healthy vegetation by subtracting the red band (2) from the near-infrared band (4) and then normalizing the ratio by dividing by the sum of the two bands.

\[ \text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}} \]

This ratio is highest (light in colour in the image below) where chlorophyll absorption is highest and low values (darker pixels) will tend to indicate bare soils, water bodies or built up areas (Mather, 1987, Drury, 1990, Barett, 1992).

![Figure 10. The Normalized Difference Vegetation Index](image)
4.2 INEGI Maps

A series of hard-copy 1:250000 scale maps of various themes were obtained through the Instituto Nacional De Estadistica Geografia Y Informatica (INEGI) in Mexico. These maps include:

- soils
- vegetation and land use
- topography
- isotherms
- isohyets
- roads and human settlements
- surface drainage

![Digitized soils map of the region with major and minor soil types represented by different colours. The study area is enclosed by the box.](image)

These map sheets were digitized in a Geographical projection (latitude/longitude) and then reprojected into a Universal Transverse Mercator (UTM) projection. Of particular
importance was the soils map. In addition to identifying the major and minor soil type(s), attributes were provided from soil pit sampling. Among them, some that were critical in reclassifying the soils map for model inputs were: ‘A’ horizon depth, rooting depth, texture (% clay, sand, silt) and organic content. The vegetation and land use maps were useful in identifying the primary species types for forests in the study area and for corroborating information in the literature as to which activities were taking place where. It was also used to assist in interpreting the MSS images and also segmenting the NDVI image for reclassification.

The isohyet data came as two maps: one for each of the wet and dry seasons. Once these lines of constant precipitation were digitized as vectors, they were rasterized using a minimum curvature operation. This was preferable to other techniques in that it provided the smoothest surface when interpolating between data points. Once raster layers were created for both the wet and dry season, they were added together to produce a raster map representing annual precipitation. This data layer was used in the model, however it was also important in comparing precipitation totals with the climate station data sets. The climate station data was used in order to determine the number of rain days per year, and because the climate stations were a few kilometers outside the study area, a close correlation was necessary in order to confidently use the data for that model input.

4.3 Precipitation Data

Precipitation data for Chiapas was acquired through Karen O’Brien who had been doing climate change research in the Selva Lacandon for the North American Commission on
Environmental Cooperation. She had obtained the data sets through the Servicio Meteorologico Nationale de Mexico. The data sets included daily precipitation data dating back to 1969 for two climate stations and monthly dating back as far for an additional two. As indicated, the climate stations are not in the actual study area but are sufficiently close as to provide important details not found in the digitized annual average total isohyet map (see Appendix B for annual rainfall data from climate stations).

Figure 12. 3-Dimensional rendering of rainfall in the study area after adding the May-Oct and the Nov-April isohyet maps together. The perspective is from the NE corner of the data set looking to the SW. Rainfall tends to increase as elevation decreases.

4.4 Other data

Digital data sets of Mexico and municipal boundaries for the state of Chiapas were obtained through Carlos Rodriguez, a GIS analyst with GTI, Inc. in Guadalajara, Mexico. A compact disc cataloguing the 1994 Mexican census was obtained through INEGI, the
national mapping agency in Mexico, but was not used directly as it the bulk of data referred to cities only.
5.0 Deriving the vegetative cover inputs

De Jong's SEMMED uses an exponential model to calculate coefficients which correlate his field data for vegetative cover to spectral indices of Landsat TM images. These coefficients are in turn used in calculations of interception, P, crop cover, C, and the ratio of actual to potential evapotranspiration, Et/Em. Without field data for land cover, a different approach had to be taken. From the MSS images, there were three main approaches evaluated in order to derive the vegetative cover model inputs. They were: supervised image classification, unsupervised image classification and NDVI segmentation.

5.1 Supervised Classification

A supervised classification of an image involves defining groups of image pixels for which the land cover class is known. This establishes patterns, or a range of spectral signatures, for a particular land cover class. This is also known as 'training' the software to recognize that pattern, so that it may then assign other image pixels to that land cover class. In the Erdas Imagine software used for this project, there are several ways to establish the patterns the software will use to create classes. One is to use the mouse to enclose a group of pixels known to be in the same class with the mouse. Another is to define regions representing classes in a scatterplot image representing pixel locations in 2-dimensions, or two bands.
The method used here was to select a pixel in the centre of a group of pixels that are either known or likely to be in the same land cover class, and have the software 'grow' the area until impeded by pixels that are spectrally different. The searching algorithm used determines whether adjacent pixels will become part of the cluster, based on a user-defined Euclidean spectral distance. A pixel joins the cluster if the digital numbers of all bands being examined are within the specified distance (ERDAS Field Guide, 1994). This is a cluster-forming operation, which can be carried out several times per class. When a sufficient number of clusters have been formed to define a class, the statistical signature of the class is calculated and saved and used as the basis for assigning pixels to the class during classification.

The inherent drawback to a supervised classification for this project is a lack of ground truthing, or the ability to definitively assign a pixel to a land cover class based on some a priori knowledge. First, the only other data that could be referred to are hardcopy maps at too small a scale for this type of function. Secondly, because land cover may change often, we would need knowledge of what land cover class existed in a particular location not only at the time of year when the satellite image was captured, but also the particular year that it was captured.

Therefore, running the supervised classification became an exercise whereby land cover classes were defined through a combination of general knowledge of the landscape, a visual survey of patterns in the image and anecdotal evidence from the literature, such as descriptions of land use patterns. Another drawback to this method, was that pixels
which are not spectrally proximate to any defined class will either not be grouped (the model inputs depend on all pixels being assigned a class), or be lumped into a class, thus making the assignment of these pixels particularly arbitrary.

5.2 Unsupervised Classification

Unsupervised classification is more automated and relies on the user for only for a few parameters which will be used by the computer to group pixels based on statistical patterns. Each image pixel is clustered with other pixels having similar spectral properties. The number of classes is determined by the user. I began with fifty clusters, or classes, even though I planned to narrow it down to only three eventually\(^8\). Creating a much larger number of clusters ensures that all pixels in a category are very spectrally proximate and allows the user to interactively merge clusters into land cover classes, rather than letting the software.

After the image had been divided into 50 clusters, a visual check on each was made by assigning a colour to a class and turning it on, or making it opaque, so that pixels in that class were overlain in the assigned colour on the image being classified (see Figure 14).

In carrying out a visual assessment like this, each of the 50 clusters were merged into one of three land cover classes: bare/built, grassland/agriculture and forest. Ultimately, the same drawback is encountered here as during a supervised classification: the operator has to make judgments with incomplete ground truth knowledge. In this case, though, the user is making the decision on each cluster, whereas the supervised classification will assign pixels to their closest category.
Figure 13. The attribute editor is used to colour and overlay clusters in the process of assigning clusters to land cover classes.

5.3 NDVI Segmentation

The NDVI image values range from -1 to +1. In order to make these values easier to work with, a 100(pixel_value + 1) transformation was done, thus modifying the data value range to 0 to 200. The image was then reclassified into the same three categories as above: 1) forest, 2) agriculture/grassland, 3) bare/built up. The methodology used was quite similar to the cluster merging described in the unsupervised classification section, except instead of examining similarities across three bands of information, only the NDVI values were looked at. I worked on the assumption that bare or built areas would have the lowest values, grassland and agriculture the next lowest and forest the highest. The issue became determining the NDVI values where the classes should be
segmented. This process involved lowering the low end of the forest class until pixels which do not appear to represent forest began to enter the class. This was a fairly arbitrary process that, without ground truth data, was supported only by visual analysis and a general idea of the landscape from viewing aerial photographs and reviewing soil, vegetation and land use/land cover maps. For the bare/built class, the upper limit of the class was increased until some pixels which displayed evidence of vegetation began to enter the class. With some minor modifications to the limits the grass/agriculture class was what was left in the middle.

Figure 14. Land cover classification derived by segmenting the NDVI image. The precipitation interception (P), crop cover (C) and ratio of real to potential evapotranspiration (Et/Eo) maps look the same and have values based on the reclassification of this product to suit the particular input. These classes were further merged by combining built and eroded and also grassland with agriculture.
5.4 Further Processing and Observations

In deriving the vegetative inputs to the model, the unsupervised classification and the NDVI segmentation techniques were both used. After land classes were defined and the data layer imported into the GIS, a cross reference was done against the slope map in an effort to improve the class designations. Two rules were defined based on information from literature and a survey of aerial photographs for the general area: 1) forest will not exist on slopes of less than 10°, and 2) agriculture will not take place on slopes greater than 35°. Where there were pixels classified as forest on slopes of less than 10°, the classification was changed to agriculture/grassland and vice versa for slopes greater than 35°.10

In the 1992 image, there were a few small areas of cloud cover along the eastern perimeter of the study area. These areas had extremely high reflectance and were consequently classified as built/bare. Using the seeding method to capture spectrally similar pixels described in the supervised classification section above, the clouded areas were isolated. These pixels were saved into another layer and exported to the GIS where they were reclassified to represent either forest of agriculture, depending on what most of the surrounding pixels were. A COVER operation was then run whereby the classified map layer was overlain with the cloud map and the new pixel values assumed from the reclassified cloud cover map.
As stated, both the NDVI and the unsupervised classification methods were used and final erosion prediction made using vegetative inputs derived from both classification schemes. The degree to which the final erosivity predictions varied will be discussed later. Below is a table which indicates the number of pixels and the overall percentage that were assigned to each land cover class. In addition to using the two classification methodologies, two different images were used: the MSS image from 1980 and the one from 1992. The purpose for this was to compare final erosion results using different methods in order to assess the impact differences in classification methods would have on the final result.

![Landcover classes by image date and classification method]

**Figure 15.** Comparison of land cover classes assigned for images from 1980, 1992 from the unsupervised classifications and segmentation of the NDVI.

Although there are some observations which could be made from surveying the results shown in the table, there is risk in doing so. For example, the image capture date of the
1980 image was November 17, the end of the rainy season, the peak of healthy biomass and chlorophyll absorption, and just after harvest time for maize. In segmenting the NDVI and also in assigning clusters to land cover classes, the image from 1980 was difficult to work with because it was captured at the peak of healthy vegetation which happens to also be just after harvest. In the areas of the image where crops are interspersed with bare and eroded land differentiating them was problematic due to the spatial resolution of image capture. Additionally, unharvested crops and other non-forest vegetation are at a peak in the image, making them more difficult to distinguish from forest.

The image capture date of the 1992 image was August 27th, well into the growing season, but two months before peak chlorophyll absorption and harvesting. The phenological differences between the two image capture dates is substantial, as is the difference in reflectance values for cropland before and after harvest. Therefore, in classifying the images with different capture months, the same criteria could not be used for both in order to surmise land cover classes. As such, to look at these classifications and conclude that there was more forest or more bare, eroded areas in 1980 than 1992 is to read more information into the data than it reveals. Comparing results from the two classification methods will either increase our confidence in results or underline the sensitivity to classification method and, therefore, vegetative inputs C, P and Et/Eo in the final results.
An additional consideration in deriving these inputs was that the image upon which the values are based is a snapshot in time, whereas the model is operating on an annual basis. The image is useful, then, in determining which land cover class exists at any location in the study area, but the model parameters must be adjusted to account for which phenological changes will take place at that location over the course of a year. Also of importance in making these assessments is the season and associated level of precipitation.

Figure 16. Maize cultivation, pasture, eroded slopes and shrub growth all are shown here in close proximity. Land cover class crossover is inevitable when attempting to classify landscapes such as this with such high spatial variability (from Collier, Fields of the Tzotzil, 1975).
For example, a field may have a crop cover (C) value of 1.0 after burning and just before the planting of the maize and beans, 0.2 during the rainy season when the crops are being cultivated, and a value of 0.4 after harvesting takes place and the stalks are left to rot. Not only should these three phases be considered rather than simply the stage that is evident in the image, but also the proportion of annual precipitation that is concurrent with each phase. Since soil erosion is a function of precipitation (we are not evaluating wind erosion here), a vulnerable state of land cover is not as much of a liability if it occurs when there is very little rainfall. Table 3 demonstrates such a calculation.

<table>
<thead>
<tr>
<th>Months</th>
<th>Land Use</th>
<th>C Value</th>
<th>% An. Precip</th>
<th>Weighted C Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan - Mar</td>
<td>scrub</td>
<td>.004</td>
<td>7</td>
<td>.00028</td>
</tr>
<tr>
<td>April - May</td>
<td>cleared</td>
<td>0.9</td>
<td>12</td>
<td>.108</td>
</tr>
<tr>
<td>June - Dec</td>
<td>maize</td>
<td>0.2</td>
<td>81</td>
<td>.16</td>
</tr>
<tr>
<td>Annual total</td>
<td></td>
<td></td>
<td></td>
<td>.26828</td>
</tr>
</tbody>
</table>

Table 2. Weighted calculation for vegetation related model parameters.

Due to the nature of the terrain and the spatial resolution of the images, I expect there to be some confusion between these classes. The Pathfinder images have been resampled from 59m x 79m to a 60m x 60m spatial resolution. The Chiapan highlands, however, display many disparate land uses occurring in close proximity to one another. The forested slopes that are too steep to support agriculture and readily discernible from the MSS image, as are valley bottoms that have given way to grassland and pasture. However, there are difficulties in identifying and distinguishing between crops, fields in
fallow and naturally occurring scrub growth. This is especially difficult where these practices border on forest or grassland (Drury, p. 286; Lo, p. 194). Finally, the task is further complicated in that several land cover patterns often exist in close proximity to each other in this often highly heterogeneous landscape.

Figure 17. Forested slopes and sparsely vegetated areas as seen here are discernible from the satellite images (from Collier, Fields of the Tzotzil, 1975).

There were three main difficulties encountered with this methodology. Firstly, field data would have been useful to isolate spectral signatures of various land cover classes such that supervised classifications would have more authority. Secondly, the spatial resolution of MSS data and the often close proximity of divergent land uses in the study area meant that reflectance values in many pixels were a product of not one, but two or
three land use classes on the ground. Finally, the differences in image capture dates (i.e. the month and it's relationship to the growing season) makes it difficult to establish baseline thresholds for NDVI values or digital number values in MSS bands such that comparisons can be made between different image years.
6.0 Spatial Data Processing

6.1 Building the Spatial Database

Spatial data processing was done using three software packages: Genamap, Erdas Imagine and GRASS. Genamap is a Unix based GIS package that has both raster modeling and vector analytical capabilities. The vector-based tools were used in data entry (digitizing), polygon formation, tagging and reclassification operations. Vector maps were then converted to raster maps using a nearest neighbour or minimum curvature interpolation depending on the type of data represented in the original vector map. Genamap was also used in most of the raster modeling operations. Raster maps were made representing each model input. For example an input to the model is the typical intensity of erosive rain in millimeters per hour. This was determined to be 25mm/hr\(^{11}\) and so a raster map of the study area was created where every pixel has the value 25. When an equation calls for this parameter to be multiplied by some other parameter, a new map is created as the product of the maps being multiplied and each pixel is the product of the spatially coincident pixels.

Genamap supports two types of raster data: discrete and continuous. In a discrete raster map, each cell or pixel has one value and that value represents the entire cell. It can be numeric, and represent something like the soil detachability index for that pixel, or it may be categorical and represent a soil type for example. Pixels in continuous raster maps have numeric values and may represent elevation or rainfall for example. Although each pixel has a value assigned to its centroid, the areas between pixel centroids have
values that are interpolated based on their proximity to values of nearby cells (as per the interpolation method of choice) and are thus considered to be continuous surfaces. Most of the raster modeling was done using discrete raster maps that were created by rasterizing vector maps or as products of mapping functions.

ERDAS Imagine software was used for the image processing component. Five Landsat scenes for Chiapas were obtained in a generic binary format on 8mm Exobyte tapes. Once read and loaded the most appropriate scene was selected and a study area defined. The study area was clipped out of each MSS scene. The image processing capabilities were utilized to carry out a supervised classification of the images and then an unsupervised classification. This created three land cover classes: forest, agriculture/grass/pasture and bare/built up. A similar process was carried out by exporting the NDVI image and creating classes by segmenting that image in the GIS. Imagine was also used in attempts to assess land cover change over time by comparing the different images, comparing the time-series classifications and subtracting the later NDVI from earlier ones in order to assess change.

GRASS, a raster-based GIS on the Unix platform, was also used. It was used for it’s hydrological routing capabilities in the water phase of the model. The DEM and overland flow raster maps were imported from Genamap and used in the r.watershed function which routes the flow based on elevation characteristics of the terrain. It creates
as a product a map for which each pixel represents the accumulated flow of water, based on which up-stream pixels drain into it.

6.2 Determining Splash Detachment and Transport Capacity

![Flow chart](image)

**Figure 18. Flow chart depicting determination of splash detachment.**

To estimate soil loss due to erosion, the model requires the calculation of the amount of soil displaced by splash detachment and its comparison with the transport capacity of the same area (or pixel in this case). Both sides of this comparison, splash detachment and transport capacity, have a water and a sediment phase. However, the discussion here focuses not on those two phases but rather on the methodology for arriving at the two erosion estimates: splash detachment and transport capacity. All equations are from Morgan, Morgan and Finney (1984), although the C, P and Et/Eo inputs were modified by de Jong (1994) to incorporate the use of satellite images and again by this study, as described in the previous section. A further modification was made to equation 5, the
transport capacity equation, and is explained in section 6.3. The schematic for arriving at the value for the amount of splash detachment is shown above in Figure 18. The calculations for the required inputs are as follows:

(1) \[ E = R \times (1.9 \times 8.7 \log I) \]

where \( E \): Kinetic energy of rainfall (J/m)
\( R \): Annual rainfall (mm)
\( I \): Typical value for intensity of erosive rainfall (mm/h)

\( R \) 1:250000 INEGI isohyet maps were digitized to create a vector data layer representing lines of equal precipitation. This map was then converted into raster format by interpolating values for each pixel from the lines of constant precipitation using a minimum curvature algorithm.

\( I \) This value was not obtainable from the Chiapan climate station data as it is not something that generally gets measured. Morgan, Morgan and Finney (1984) and Morgan (1986) both mention some typical values for different climate types and this information was used to arrive at a figure of 25mm/hr for highland Chiapas.\(^{12}\)

(2) \[ F = 0.001 \times K \times (E \times \exp(-0.05\times P)) \]

where \( F \): Splash detachment (kg/m\(^2\))
\( E \): Kinetic energy of rainfall [from equation 1 above]
\( K \): Soil detachability index (g/J)
\( P \): Interception map (%)

\( K \) This value is related to the soil particle size and comes from experimental work by Quansah (1981). Correlation tables between the Quansah K-values and soil textural classes were provided by de Jong (1994) and Morgan (1986). The soil map was reclassified based on textural class attributes to provide the K map.
Interception values were determined by reclassifying the land cover/landuse map which was in turn created by classifying the MSS images. The lookup table between land cover classes and P values was created referencing Morgan (1986).

Figure 19. Results of splash detachment (F) calculations.

Distributed transport capacity was arrived at by following the sequence outlined in Figure 20 below.

The overland flow per pixel is determined with the following equation:

\[ Q = R \cdot \exp(-Rc/Ro) \]

where

- \( Q \): Volume of overland flow (mm)
- \( R \): Annual rainfall (mm) [from isoyhet maps]
- \( Rc \): Critical value of moisture storage [from equation 4 below]
- \( Ro \): Annual rainfall divided by the number of rain days per year
Ro Using the "rain day" definition from de Jong (1994), this value was calculated by averaging the number of rain days for the particular year (synched with the image year) and averaged over four climate stations. A review of the climate station data confirmed that areas with higher annual rainfall also have a higher number of rain days. A graded scale was created for this input where areas of higher precipitation had a higher number of rain days.

Figure 20. Flow chart of data processing steps to arrive at the distributed transport capacity.

\( \text{(4) } \quad R_c = 1000 \times MS \times BD \times RD \times (\frac{Et}{Eo})^{1/2} \)

where
- \( R_c \): Critical value of moisture storage (mm)
- \( MS \): Soil moisture content at field capacity (w/w)
- \( BD \): Bulk density of top soil (g/cm³)
- \( RD \): Top soil rooting depth (m)
- \( \frac{Et}{Eo} \): Ratio of actual to potential evapotranspiration
MS  Soil moisture content at field capacity was arrived at by consulting Morgan (1986) and reclassifying the soil map.

BD  Bulk density was arrived at in the same manner.

Et/Eo  Values for the actual to potential evapotranspiration ratio were arrived at from norms identified in the literature (Morgan, 1986, Atlas Nacional de Mexico, 1995, Kirkby, 1978) and assigned by reclassifying the land cover classes determined by the image classification process described previously. Values for forest were 0.95 and bare/built were assigned 0.05. Typical values for maize are 0.6-0.7 and values for grassland are 0.8-0.95. However, fields which contain maize for half the year will have extremely low values for other half as they will be bare for some of the time and, depending on soil fertility and altitude, may have vegetation for some of that period. The value that was used to try and cover these various scenarios was 0.7.

RD  The rooting depth coverage was also reclassified from the soils map with the values taken directly from soil polygon attribute information. However, as rooting depth is a function of ground cover and will, therefore, change with time, these values were not necessarily accurate. Higher rooting depth values (as would indicate forest, for example) as taken from the soils map, would overestimate the soils capacity to absorb water if the vegetation had since been removed. The land cover map, created by classifying the satellite images, was reclassified to represent typical rooting depths by vegetation type. Then, the rooting depth from the soils map and the reclassified land cover map were overlain and the minimum value taken as the rooting depth for use in this equation. Figure 21 shows the result of this process. It is easy to see where the large soil polygons have pixels that were reclassified by this operation due to the splattered effect they create. Where the values from the soil map were altered it is a case where the area was considered bare ground from the satellite image, yet had a high rooting depth assigned from the soils map.
This Q value determined from equation 3 above indicates the volume of overland flow for each individual pixel. In order to account for where this flow goes and which other pixels may flow into a given pixel, a hydrological routing algorithm was used. The overland flow map and DEM were imported into GRASS and run through a watershed routine. This routing algorithm produces an accumulated flow map after having run the flow input map over the DEM. This “distributed” result map was brought back to the GIS and used in subsequent calculations.

![Diagram with legend](image)

**Figure 21.** Results of taking the minimum value from the soils map rooting depth values and the reclassified land cover map values.

The pixel values generated by the hydrological routing routine were sufficiently large that the Genamap import function, which only handles two byte data, failed to read the file. A translator was written which processed an ASCII-dump of the GRASS raster file
and assigned the values to the appropriate row and column in a Genamap raster map. Though processing is extremely slow, this translator worked where innumerable other attempts had failed to get the results of the hydrological routing algorithm back into the GIS. The accumulated flow map was then used in the following equation to determine the transport capacity, where transport capacity refers to the flow or transport of soil particles by runoff (de Jong, 1994, p.79). This equation has been slightly modified in order to account for discrepancy between observed flow values from the INEGI surface flow map and calculated overland flow values. The rationale for the modifications is discussed in the following section.

\[
G = C \times ( \frac{3}{5} \times Q)^2 \times \sin(S) \times 0.001
\]

where

- **G**: Transport capacity (kg/m²)
- **C**: Crop cover (USLE-C)
- **Q**: Accumulated overland flow volume (mm) [from 3 above]
- **S**: Slope in degrees

**C** is derived from the classified MSS image. Values for each land cover class were assigned values based on Morgan's lookup table (1986).

**S** The percent slope was derived from the DEM with a built-in GIS function and then converted to degrees using the following equation:

\[
slope\_degrees = \text{atan} \left( \frac{\text{slope\_percent}}{100} \right)
\]

The transport capacity map was then compared with the splash detachment map and, for each pixel, the minimum value of the two was taken as the quantity of soil lost to erosion per year (kg/m²). The minimum value is taken, because the lower of the two is considered to be the limiting factor in the erosion process (see Figures 23-27).
These procedures were carried out with the vegetative inputs derived from the 1980 and the 1992 image separately and also by the unsupervised classification and the NDVI segmentation technique. This yields four separate results and a fifth was created using de Jong's technique of subtracting the saturated infiltration capacity with the unsupervised classification of the 1992 image. They are discussed in the results section.

![Kg/m²](image)

**Figure 22.** The results of the transport capacity calculations.

### 6.3 Calibrating the Overland Flow Values

The result from equation 3 represents overland flow per pixel (Q). A correction was made to these results in order to account for infiltration of runoff which is high in the study area. According to the surface drainage map (SD, digitized from INEGI source map at 1:250000), most of the study area is in a 5-10% and 10-15% range where the values represent the percentage of precipitation which drains over the surface. These
surface drainage map values were compared with the overland flow values using the following procedure.

The DEM was input to a hydrological routing algorithm in GRASS which output a map in which each pixel value represented the number of upstream pixels which flow into it.\textsuperscript{15} The cumulative flow map (as described in the previous section) was then divided by this unit-value flow map to determine the average flow per upstream pixel (AF). Annual precipitation (R) was divided by this result so that the average flow per upstream pixel, Z, was represented as a percentage of annual precipitation.

\[ Z = \frac{AF}{R} \times 100 \]

These values were then compared with the surface drainage map (SD), which represents observed flows expressed as percentages of annual precipitation, and were found to be higher. In order to establish a calibration coefficient, a weighted average was calculated for both maps and the surface drainage map average (16.9\%) was determined to be 3/5 the value of the Z map average (28.2\%).

\[ \text{Calibration coefficient} = \frac{\text{SDavg}}{\text{Zavg}} = \frac{3}{5} \]

This coefficient was then applied to the overland flow value, Q, in the calculation of transport capacity, G, in equation 5 above.

In SEMMED, de Jong proposes the subtraction of the saturated conductivity of the soil map from the distributed overland flow in order to correct the model for the same
purpose. For several reasons, this is not ideal in this case. First, the calculations for flow are on an annual basis, whereas the saturated conductivity that de Jong subtracts is a daily maximum by soil type. Other unknowns detract from confidence in this method, such as the high likelihood during the rainy season that soils become waterlogged and do not exhibit their full capacity for absorption. In order to compare results, I have used both methods here, and the differences will be discussed in the results section. Saturated conductivity values by soil type were obtained (Dingman, 1994, p.222) and the soil map reclassified to reflect them. This was then subtracted from the overland flow map before transport capacity was calculated.
7.0 Results and Observations

7.1 Erosion Hazard Maps

Figures 23 through 27 show the results of the discussed data processing steps. Predicted rates of erosion range from just above 0 to 10kg/m². Although results vary due to different image years and classification techniques, the proportions across erosion classes are similar (see Figure 28). Depending on image classification technique used, between 40 and 55% of pixels are predicted to show erosion levels of less than 1kg/m²/yr and about 14% show predicted levels of greater than 4kg/m²/yr. Discernible in the erosion map are the major soil type polygon boundaries, indicating the model is sensitive to certain values obtained from soil attributes.

![Erosion Hazard Map](image)

Kg/m²

- □ 0-1
- □ 1-2
- □ 2-3
- □ 3-4
- ▪ 4-5
- ■ 5-6
- ■ 6-10
- ■ >10

Figure 23. Predicted erosion results using land cover classification done using the 1980 image and the unsupervised classification.
Figure 24. Predicted erosion results using land cover classification done using the 1980 image and the NDVI.

Figure 25. Predicted erosion results using land cover classification done using the 1992 image and the unsupervised classification.
Figure 26. Predicted erosion results using land cover classification done using the 1992 image and the NDVI segmentation.

Figure 27. Predicted erosion results using land cover classification done using the 1992 image and the unsupervised classification. Saturated infiltration capacity was subtracted from overland flow rather than using the surface drainage map to mitigate those values.
Shown in Figure 27 are final predicted erosion results when de Jong's method of subtracting the saturated infiltration capacity from the distributed overland flow is used. These values are slightly lower, but follow the same pattern, i.e. there are differences of magnitude, but the relative differences within the different results shown are similar. Although he used a different method to account for infiltration, he was likely attempting to do the same thing: bring predicted values into line with observed values.

Figure 28. Distribution of predicted erosion values by image date and classification method.

As shown in Figure 28, the predicted erosion values tend to be higher when the 1992 image was used, although this has to do with the judgments made during classification and not necessarily changes in land cover over that period of time. Land cover assignments clearly impact these results, as can be seen by comparing this graph with
Figure 15 which displays class assignments. The two unsupervised classifications had the highest proportions of built/bare pixels and, correspondingly, show the highest proportions of predicted values above 6kg/m$^2$/yr. At the low end, the 1980 unsupervised classification and the 1980 NDVI segmentation had the two largest forest classes, and they have the largest number of pixels between 0 and 1kg/m$^2$/yr.

![Figure 15](image)

*Figure 29. A difference image between different classification methods and the predicted erosion results from the 1992 image.*

Figure 29 depicts the difference between the predicted erosion results for 1992 which were derived from the unsupervised classification and those obtained when the NDVI classification was used. The white pixels, which represent about 85% of the total, were the same predicted erosion values in both scenarios. The blue pixels indicate where the unsupervised classification yielded higher predicted rates of erosion (cyan show a difference of 1-2 kg/m$^2$, navy are 2-4 kg/m$^2$) and are indicative primarily of areas where
pixels were assigned to forest by the NDVI classification and to agriculture in the unsupervised classification. Pixels which were classified as agriculture by the NDVI process and as built/bare by the unsupervised classification are shown in orange and red (orange is a predicted difference of 1-2 kg/m² and red is 3-4 kg/m²) and are more numerous.

The variance in predicted values for the same year result from the same pixel having been classified differently in the two methods used. This is as a result of some of the uncertainties already discussed with the classification process. For example, there is not necessarily a large difference between the NDVI value for senescent crops and a forested area, so in using the NDVI it is not surprising there was some cross over between classes. This highlights the differences in the approaches used and also show how important the classification process was to the final results. Most pixels were classified the same in both procedures, but where different assignments were made, the predicted erosion results varied, sometimes by very significant levels of more than 2 kg/m²/yr.

Given the uncertainty of accurate pixel classification, it is interesting to look at the more generalized patterns in the predicted erosion levels. Figure 30 shows the results from the 1982 image re-sampled using bi-linear interpolation to a 240 by 240m resolution from the original 60 by 60m (predicted erosion rates follow the same colour scheme as in Figures 24-27). Considering the difficulties in land cover classification for specific pixels, these smoothed results are useful in that they show the general patterns of erosion
hazard at a resolution we are more confident with. This is perhaps a more appropriate resolution to view the results as it gives a more general idea of erosion hazard patterns across the study area rather than specific predictions at exact locations. Although it is based on the same results, it directs the observer towards making more general observations for which the results are more germane.

Figure 30. A generalized representation of the 1980 results.

Figure 31 depicts which pixels in the final results were limited by transport capacity and which were limited by splash detachment (green). This was determined by dividing the erosion map by the splash detachment map and selecting from the result all pixels with a value of one. Streams and runoff channels are generally detachment limited because the transport values for these pixels were extremely high as many other upstream pixels would flow into them. The opposite situation occurs on the tops of ridges because there
is little opportunity for neighbouring pixels to contribute to cumulative flow calculations. In these cases, the results tended to be transport limited even though these areas are forested for the most part. Clearly visible in these results are soil polygon boundaries. The fairly contiguous, black areas in the above image show areas where erosion is limited by transport capacity. These conform to the highest K value, 0.4, which yield high results for splash detachment and a greater likelihood that the prediction will be limited by transport capacity.

![Image](image.png)

Figure 31. Areas in green show where the final results were limited by splash detachment (F) and black shows transport capacity (G) limited results.
The model's basis for calculating transport capacity probably leads to higher results than is should. The flow per pixel is routed over the DEM to determine the distributed overland flow, and this value is subsequently influential in determining the transport capacity. However, no accommodation is made for the fact that some or all of the transport capacity of an area will be used up by transporting upstream sediment load. Therefore, where values are transport limited, i.e. the calculated transport capacity is taken as the predicted erosion rate, this value is high for the stated reason. Also, where results are detachment limited, it is possible that with a transport capacity which reflected sediment load, the result would be transport limited.

7.2 Model Sensitivities

In an effort to ascertain which values or combination of values were most influential in the final results, sensitivity tests were run. The Morgan, Morgan and Finney approach was to modify individual input values and monitor changes in predicted erosion to assess the sensitivity of the model to the input. Since the equations used here are the same, so would be the results. The approach here has been to work backwards from the result maps to look for telling patterns in individual inputs. The final erosion maps were divided into categories which match the legend shown in the above figures. A SCORE operation was done using the GIS to calculate what the average value was in various input maps for each category in the erosion map.

This table shows general data trends that are expected. Precipitation interception, for example, interferes with the erosive impact of rainfall, and the higher erosion categories
all show the lowest value of precipitation interception. A very similar pattern is displayed for C, the crop cover input, which is also derived from the classified MSS images (except high values are seen at high erosion levels because C values increase towards one as vegetative cover decreases). The third input derived from the MSS images, EtEo shows fairly even values across the ranges, indicating the model is not very sensitive to changes in this input.

<table>
<thead>
<tr>
<th>Model Input and Range</th>
<th>0-1</th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
<th>6-10</th>
<th>&gt; 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>P (10-30)</td>
<td>27.32</td>
<td>26.42</td>
<td>25.70</td>
<td>24.48</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
</tr>
<tr>
<td>K (0.2-0.4)</td>
<td>0.36</td>
<td>0.30</td>
<td>0.25</td>
<td>0.31</td>
<td>0.39</td>
<td>0.22</td>
<td>0.33</td>
<td>0.40</td>
</tr>
<tr>
<td>C (0.002-0.7)</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.14</td>
<td>0.17</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>R (1496-1803)</td>
<td>1752</td>
<td>1748</td>
<td>1757</td>
<td>1726</td>
<td>1731</td>
<td>1777</td>
<td>1701</td>
<td>1766</td>
</tr>
<tr>
<td>RD (0.05-0.28)</td>
<td>0.19</td>
<td>0.14</td>
<td>0.14</td>
<td>0.16</td>
<td>0.13</td>
<td>0.10</td>
<td>0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>EtEo (0.7-0.95)</td>
<td>0.78</td>
<td>0.73</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td>0.77</td>
<td>0.68</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Table 3. Average values from selected model input parameters by erosion range (columns).

As erosion values increase, so does the slope of the terrain, however, slope drops again at the highest erosion values of > 6 Kg/m². This could be explained by the fact that only about 6% of pixels fall into that range and most are built or bare areas with no vegetative cover, making them susceptible to erosion. Another explanation is that where slope is very high, so are transport capacity values and so results are splash detachment limited. Rooting depth averages fluctuate, but do show a general trend of being high where erosion is low and low where erosion is high.
The SCORE method shows average values by erosion category of some other model input. It does not indicate the distribution of values that contribute to the average value. To do this, a single rate of erosion range is selected, such as $<1$ or $>4$ kg/m$^2$ and all pixels in the result map in the range are extracted and assigned a value of 1. This intermediate map was then multiplied by the input of interest, say K, and the result is a map of original K values, but where only pixels which had erosion rates of $<1$ have values. Figure 32 below demonstrates this process.

<table>
<thead>
<tr>
<th>Erosion values</th>
<th>Isolate $\geq 4$</th>
<th>Slope in degrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 5 4</td>
<td>1 1 1</td>
<td>9 12 11</td>
</tr>
<tr>
<td>2 1 1</td>
<td>0 0 0</td>
<td>5 6 4</td>
</tr>
<tr>
<td>6 3 5</td>
<td>1 0 1</td>
<td>8 3 18</td>
</tr>
<tr>
<td>Slope where erosion $\geq 4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 12 11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By isolating a range of erosion values, in this case those greater than 4 kg/m$^2$/yr, the corresponding slope values are obtained. This allows examining the distribution values from a particular model input that are associated with a particular category of erosion values.

Figure 32. Methodology for determining distribution of pixels from a model input parameter for a specific range of predicted erosion.

Sensitivity can be looked at not only by final results, but based on what the limiting factor in the final calculation was also. The impact of a particular input can be more
telling when looked at in this manner. For example, we see from Table 3 that the average precipitation interception value (P) tended to drop as predicted erosion increased. This input is only used in the detachment equation, and when only results which were detachment limited are looked at, we find the average P value of 29.59. This indicates that almost all pixels where final erosion predictions were detachment limited had P values of 30, the highest in the range.

In Table 3, the range of values for rooting depth are ambiguous. However, rooting depth is only an input to the transport capacity calculations. By isolating the pixels from a result map which were transport limited, and which had high values for predicted erosion (using the methodology described in Figure 32), a different view of the data is obtained (see Figure 33). Where transport limits the erosion prediction, RD is quite influential, as can be seen by comparing the distribution of pixels against that of the original RD map. One reason for this is that the range of values is relatively large (0.05-0.28), and for this reason it has more impact on results than other inputs from the critical value of moisture storage calculation (Equation 3) such as bulk density.
Figure 33. Lower rooting depth values contribute significantly to high predicted erosion rates.

When the same calculation is made with slope, the results show that slope is not a significant factor in the final results, as the distribution of slope values across categories is almost identical to the distribution of the original slope map. It is evident that the model is sensitive to the splash detachment index, K, because where results are detachment limited, the distribution of K values is significantly different than for the entire area. In the original map, over 50% of pixels are assigned a value of 0.4 and about 10% are assigned 0.2. In detachment limited results, there are only half as many at 0.4 and three times as many pixels have a value of 0.2. This indicates that where K is high, so is the predicted splash detachment and so the results tend to be limited by transport capacity (where K=0.4, almost 90% of pixels have transport limited estimates).
Figure 33. Distribution of K values shifts when predicted erosion results are limited by splash detachment.

These results are generally consistent with the sensitivity testing carried out by Morgan, Morgan and Finney (1982) and de Jong (1994). Where splash detachment limits erosion, rainfall (R) is the input to which the model is most sensitive. Precipitation interception (P) and the soil detachability index (K) are both inputs to which the model is sensitive, and by testing as described above it is shown that sandy loam soils and areas which have low P values constitute high erosion risks.

7.3 Model and Methodology Shortcomings

This model and the given data sets serve to provide a general index of soil erosion hazard in the study area. This project necessitated modifications to the SEMMED model in order to work with particular conditions associated with the study area and as a result of different data sets. Substantial enhancements or other modifications to the model were beyond the scope of the project. However, there are a number of issues that have arisen.
and which warrant comment. Some speak to shortcomings in the input data and others to limitations in the model.

1) The soil polygons were digitized from 250,000 scale maps and are up to 100km². There is likely to be a high variability in soil properties over such a large area and these are not reflected. Numerous model inputs were reclassification operations of the original soil polygons based on soil attribute values. Since certain of the soil attributes are extremely influential in the determination of final results, the viewing of such large mapping units as homogeneous entities must be done with some skepticism.

2) Certain soils attribute information was inferred from neighbouring soil pit attributes based on a commonality of major and minor soil type. These values may be used effectively as heuristic and generalized values for the mapping area, but cannot be considered highly accurate on a pixel by pixel basis.

3) There were also issues of scale attached to the assignment of the vegetation and land cover model inputs and interpretation of the satellite images on which they are based. These issues revolved around spatial resolution limitations and in discerning different land uses occurring in close proximity to one another. Initially, the spatial resolution also led to the decision to use only three land cover classes, where four or five could have been used with SPOT or TM data. Areas of mixed land cover were difficult to delineate and were grouped into general categories for this analysis.

4) Temporal resolution issues are also a factor. Change detection over time was not possible in this study, even though satellite images were obtained from three different
time periods. Although many changes occurred in the study area from 1974-1992, the phenological changes between the image capture months of August, November and February made useful change monitoring impossible. For the 1980 and 1984 images, which were captured in late November, the maize harvest had just taken place in the highlands. This made distinguishing between recently harvested fields and eroded slopes a difficult proposition. Finally, at certain times of the year, healthy crop cover will be difficult to distinguish from forest or scrub growth. This is not a limitation or difficulty inherent to this model or this study. However, it underscores the importance of selecting image capture dates (when possible) that are most suitable for the study area climate and goals of interpretation.

5) The images provide a snapshot or a single point in time data set, yet the model works on an annual basis. From a single snapshot the analyst is to derive not only what the image represents on the ground, but also what happens at that location for the rest of the year based on the initial findings.

6) The model has no accommodation for topography in its calculations of splash detachment; in other words, splash transport is ignored. In Quansah’s work (1981) upon which the K values are based, he finds that along with rainfall intensity, slope is an important determinant in splash detachment.

7) On the transport capacity side of the equation, the model doesn’t account for the upstream sediment load. By routing the overland flow per pixel, the accumulated flow is calculated. This value represents the total amount of flow passing through the pixel, however it should be recognized that the transport potential of this flow for any given
pixel may be partially or fully expended by carrying upstream sediment. A useful improvement to the model would be an attempt to account for sediment load and sediment deposition. An approach to this would be to take the splash detachment results and route them over the DEM like overland flow results. Then, on a pixel by pixel basis (in a similar manner to which the flow routing is done) use cumulative sediment flow to mitigate cumulative overland flow.

8) The model assumes that detachment is a function of the kinetic energy generated by rainfall and ignores detachment by overland flow.

9) SEMMED's accounting for saturated infiltration capacity is not highly accurate because it subtracts daily maximum values for infiltration based on soil type from annual overland flow totals (Q). This is a difficult variable to calculate for because it is a function of many other variables, such as soil type, soil particle size, porosity and the extent to which the infiltration capacity is already used at any given point in time. In a protracted rainy season like Chiapas', the infiltration capacity of the soil may be expended for long periods of time. The method used here was to calibrate the overland flow values with the use of a map of observed surface flow values. However, this data is highly aggregated (source map is 1:250000) and likely represents catch-basin averages. Nonetheless, it does in some way address the issue of high infiltration rates in the study area.

10) Noteworthy also is that, when the NDVI was used for the classifications, the results are more tuned towards amounts of healthy vegetation and chlorophyll absorption than precise land cover classes.
11) In the 1992 image there were some small areas with cloud cover. When classifying the image, these clouded areas were assigned the class most commonly found in surrounding pixels. Although these areas are small and the class that was assigned is fairly likely to be correct, this was an arbitrary measure.

12) The DEM makes generalizations as to the slope of each pixel by taking the average for the area. This may smooth out short, steep slopes where gullying can occur.
8.0 Summary

8.1 Review of Research Questions

1. What are the spatial-resolution imposed limitations of using Landsat Multi Spectral Scanner (MSS) data for the type of vegetation analysis required by this and other soil erosion models?

There are difficulties associated with classifying the MSS images for the study area into other than broad categories. Primarily, the difficulty is that land cover classes are highly variable on the ground in the study area and 60mX60m resolution will, in many instances, represent a combination of classes. This would not necessarily cause difficulty in the case of a different landscape, for example a managed agricultural one where fields are either growing crops or in fallow and the boundaries are clearly demarcated.

2. What are temporal resolution-imposed limitations of using RS data for studying land-cover change in a two season climate such as Chiapas’?

There are two central issues with respect to this question that I encountered: image classification and change detection. In classifying an image, it is essential to consider the capture date in order to properly interpret the image. Additional knowledge about what seasonal changes take place in vegetation in the image is also essential. Using this knowledge, one can construct a classification scheme of land cover classes appearing in the image, at the time of image capture. In this case, because the image was being used to derive inputs that represented annual values, further assessment was required. For each land cover class, decisions were made as to what happens to that land cover class as the season changes, and values adjusted accordingly.
Image capture date is of the essence when attempting to carry out change detection study, in particular where seasons are evident and phenology is significant. Assessing changes over time is greatly aided by images which were captured at close to the same date in their respective year. Calibrating comparisons which are done between images of different times of year is difficult and would require strong supporting data to distinguish between seasonal change and extended change over time.

3. Can MSS data be useful for assessing vegetative cover (in the way required by this study) in the absence of field data to assist in image analysis and substantiate findings?

MSS images can be used effectively to assess land cover change over time, and are used regularly in agricultural crop monitoring. Depending on the purpose and desired results, assessing vegetative cover and its change over time with MSS images may be feasible. Determining factors will be surrogate data sets and information, knowledge of the landscape, phenology, climate etc. and synchronization of image capture dates.

4. Is the SEMMED model easily adaptable easily to the Chiapan climate, given that it is designed for a Mediterranean environment? What modifications will be required?

While there were many changes in input values and minor modifications made to the model, information sources were available in order to make certain changes with confidence. Differences in input values, for example the typical intensity of rainfall, didn’t alter the model at all, only the results. The most significant changes that were
made involved the methodology for deriving the vegetative inputs to the model. This was a data issue, rather than having to do with climate. However, confidence in the values derived from the image classifications is not as high as certain other model inputs. The other significant modification was the approach to infiltration capacity. Due to high soil permeability in the study area and the availability of a surface drainage map, a different approach was taken, as described in the data processing section.

5. Can Remote Sensing and GIS be useful technologies in environment-conflict research?

In the absence of current, accurate data sets pertaining to environmental degradation and resource degradation, I believe environment-conflict is one of the many fields that will benefit from the improving accessibility of remotely sensed data sets. In addition to studies such as this one, which extracts information from the image data for use elsewhere, the satellite images provide a wealth of potential data sources themselves. Desertification, deforestation, water pollution, expanding human settlements and changes in habitat are among the many phenomenon which can be evaluated with remotely sensed data. There are three main advantages that I see to remotely sensed data for the environment-conflict researcher:

1) RS data is available covering large regions. This is relevant in that this research is generally interested in regional-scale environmental degradation and other data sets may not provide sufficient coverage.
2) RS data is current. Other data sets may be ten or twenty or more years old which is of limited use in assessing the current state of resources.

3) RS data provides the ability to monitor changes over time for a given area. This is essential for quantification of resource depletion and degradation. Other studies and data sources, which may quantify a resource accurately, are often not carried out on a regular basis. They may give good data for one time period but assessing change over time may be difficult.

As demonstrated in this study, GIS can perform complex analysis on large, regional data sets. Data sets may come digitally, from another GIS or image processing software, or be entered manually, as when hard copy maps are digitized. This study focused on the raster map mathematics and cartographic modeling functionality in the GIS. A great breadth of other analytical functions that are commonly used were not touched on here.

As with remotely sensed data, the volume and quality of digital, GIS-ready data sets is increasing rapidly, and has been for years. Digital maps of the entire world are available and can be linked to resource and environment-related databases which institutions like World Resource Institute are now making available digitally.

In addition to the wealth of information that can be obtained or created with simple GIS techniques, more sophisticated GIS and RS-based studies are available to researchers. It may not often be feasible to support an environment-conflict case-study or research initiative with resource-oriented data collection and field-work. However, the
environment-conflict researcher can benefit from being conversant in these technologies in order to take advantage of the considerable, and rapidly growing, body of environmental monitoring studies which are based in RS and GIS.

8.2 Conclusion

The modeling of soil erosion at the regional scale, as has been undertaken in this project, is greatly facilitated by remote sensing and geographic information systems. Due to the heterogeneity of the landscape and its large spatial variability, the model has to accommodate the impact of upslope water flow. The SEMMED model uses satellite imagery to derive vegetative inputs and a digital elevation model to generate overland flow. The cartographic modeling capabilities of the GIS are used extensively to combine input layers (raster maps) in the model’s various equations.

Landsat MSS images were used to derive precipitation interception, crop cover and the ratio of real to potential evapotranspiration. The lack of field data and relatively coarse spatial resolution detracted from high confidence levels in the land cover classifications made from the images, however general characteristics could be estimated. The majority of the inputs to the GIS were from digitized soil maps which were reclassified to represent attributes of the soils. Precipitation data were digitized from isohyet maps and also taken from meteorological station data sets. Overland flow per pixel values were routed over the digital elevation model using a hydrological routing algorithm to generate distributed flow values.
In general, the most important data processing step was deriving the precipitation interception and crop cover values by classifying the satellite images. Crop cover, in particular, had a large difference in values assigned to pixels depending on land cover class, thus making the classification critical to final results. Where the final erosion estimates were limited by splash detachment, annual rainfall and precipitation interception were the most influential model inputs. Where transport limited, the surface flow correction and crop cover were the most influential.

The majority of results were in the 0.5-5kg/m/yr range. However, results should not be looked at as highly accurate estimates of erosion losses in kg/m²/year but as a relative index of where conditions exist that the land is prone to soil erosion. The drawbacks to the MMF model and SEMMED are part of the trade-off that always exists between complexity or accuracy and ease of use. Morgan, Morgan and Finney remarked from the outset that a simple predictive model was their goal. More complex models are not well supported by data sets with high levels of uncertainty, and results are no more trustworthy as data uncertainties and error get propagated. This thesis provides a useful data set for the central highlands of Chiapas and a relatively simple methodology for erosion hazard mapping using accessible data sets.

The soil erosion hazard maps that have been generated indicate that due to land use patterns, soil types, rugged topography and high precipitation levels, the soils in many parts of the study area are prone to high levels of erosion. Population growth and, as far
as the Chiapan campesinos are concerned, an inadequate supply of agricultural land have led to large-scale forest clearing in order to open up new lands to agriculture. These areas are not productive for long and the swidden agriculture practiced in the study area is not sustainable.

These results show patterns of erosion hazard that are often high, and while the precise quantitative predictions for specific locations may be viewed with circumspection, the general patterns are instructive. The significant areas which exhibit moderate and severe erosion hazard are not likely to remain agriculturally productive for long and in many cases they are already pasture. This cycle of decline worsens as forested land available for clearing dwindles and the inexorable surge in population continues. Declining agricultural productivity impacts not only the income of the Mayan farmers, but the entire rural economy to which it is the central pillar. The increasingly difficult struggle for the farmers to provide basic nourishment to their families fosters the grievance process which is central to violent uprising. This is not a distant relationship, but one which is tangible and real.

Data sets and analysis such as this are critical for the political analysts who wish to evaluate resource scarcity’s contribution to economic decline in an industry or locality. This is especially the case where other relevant economic data are not recorded or available. The relationship between resource scarcity and economic decline may be fairly obvious. And it follows in certain circumstances that economic decline may
increase grievances towards the state by certain groups. However, it is crucial that evidence supporting the rate of resource degradation be more than anecdotal in order to establish environmental factors as serious contributors to the manifestation of violent conflict.

1 The project focused on water scarcity in the Nile and Jordan rivers as examples of simple scarcity, Bangladesh-Assam as a group identity conflict and the Philippines and China as examples of economic decline leading to civil strife. The Chiapas case study discussed here is also an example of where economic decline may lead to civil strife.
2 Information cited by Howard et al, 1996 and compiled from Chiapas: Habiantes de Lengua Indigena (Aguascalientes, Ags., Mexico, 1993) indicates that education and literacy levels are much lower for Mexico’s indigenous than for the population on the whole. Furthermore, the Chiapan indigenous fare worse than other of Mexico’s indigenous. More than half are illiterate and fewer than half receive any educational instruction at all. Fewer than 7% receive any schooling past the primary level.
3 See Howard, p. 5 and pp. 15-16. Also George Collier in Basta! and Benjamin in A Rich Land, A Poor People.
4 Reported as greater than 4% per annum by the 1994 Census, which translates into a population doubling time of only 17 years.
5 Cisneros did so for certain municipalities, as did the Censo Ejidal of 1994, however these works spoke only to specific areas and not to regional figures or the sparsely populated lowlands where much of the deforestation has taken place.
7 The decline in productivity is an inevitability, and one which occurs fairly quickly in this area, however there are initial increases in output associated with clearing land and increased erosion. Erosion rates of less than 1kg/m²/yr are generally sustainable, and above that the level of erosion leads to eventual degradation of fertility and decreased output.
8 This number was decided upon fairly randomly. A large number was chosen to make the clusters narrow so that merging them into the final land cover classes was a more interactive process.
9 The land use/land cover map was at a scale of 1:250 000 and has polygons representing forest, grassland and seasonal agriculture. However, the polygons are highly aggregated representations of what general class is dominant in an area. From personal experience and from reviewing aerial photographs, it is clear that the landscape is highly heterogeneous. As such the classified images are considered to be more accurate representations of ground cover classes.
10 An attempt was made to use this criteria to improve the classifications. In other words, to take the pixels which, according to this test, have been classified in error and then return to the cluster (in the case
of an unsupervised classification) and re-assign all pixels in the cluster. However, certain of the re-assigned pixels would fail the test again under the new designation. This is indicative of crossover between land cover classes (if we assume the rule to be a valid one), and I decided the best way to proceed was to simply re-assign the offending pixels rather than the entire cluster (or NDVI value).

Sources of information for this and other input parameters are discussed when data processing steps are detailed.

Morgan (1986) lists typical values of 11 mm/hr for temperate climates, 25 mm/hr for tropical climates and 30 for Mediterranean climates. The study area is in an inter-tropical convergence zone, and it is likely that the intensity of precipitation varies with the type of precipitation, which in turn differs with altitude. As such, this must be viewed as a source of uncertainty when viewing results. Sensitivity testing by Morgan and de Jong both indicate, however, that the model is not very sensitive to this input parameter.

These values are taken directly from Morgan, however they are corroborated by information in the Atlas Nacional de Mexico, and Kirkby, 1978.

Typical values for grass as given by Morgan (1986) is .10 m. In discussion with Larry Band, a value of .3 m was decided upon for forest and I have assigned .05 m for bare/built areas.

This is the same algorithm used to route surface flow and generate a distributed overland flow map. In this instance, no flow map was provided. In effect, a unit value map was routed so that the result represents the number of pixels flowing into a particular pixel, rather than the accumulation of flow values.
Bibliography


O'Brien, K. “Deforestation and Climate Change in the Selva Lacondona of Chiapas, Mexico”, Phd. dissertation, 1995, University of Pennsylvania


