Chapter 2: Exploring associations between socioeconomic variables and environmental factors: A literature review

This chapter aims to develop a set of research objectives required to fulfil the research aim in Section 1.5. To do this, Section 2.1 reviews potential associations between rural socioeconomic variables and environmental factors that have been found in the literature. Section 2.2 discusses methodological issues related to linking socioeconomic data and remotely sensed data. Section 2.3 briefly outlines the statistical analysis methods required to explore the associations between census derived socioeconomic data and environmental factors derived from remotely sensed data. The chapter concludes with a series of objectives that will be required to fulfil the research aim. Often the terms socioeconomic variables and poverty have been used interchangeably as it can be argued that poverty is partly the occurrence of a range of poor socioeconomic variables (World Bank 2000). The term environmental metric has been taken to represent any type of land use/land cover that can be detected from fine spatial resolution satellite sensor imagery.

2.1 Associations between poverty and the environment

In the past it has been cited that poverty and poor people cause environmental degradation due to short term survival taking precedence over longer term ecological protection. However, in more recent times this has been revisited and a more complex argument devised which identifies the multitude of other interactions which can also contribute to environmental degradation and poverty, such as micro and macro scale economic, political, environmental and social issues (de Sherbinin et al. 2007; Gray and Moseley 2005; Lambin et al. 2003; Giest and Lambin 2002; Scherr 2000).

Poverty is often separated into chronic poverty, where poverty persists for many years (Hulme and Shepherd 2003) and households ascending out of or descending into poverty. Often factors found to be correlated to these categories are different (Hulme and Shepherd 2003; Krishna 2007) and can also vary spatially and temporally. In most cases the determinants of chronic poverty, descent into poverty or ascent out of poverty vary. In Bangladesh, Sen (2003) found that; remote locations, unfavourable agricultural conditions, limited access to transport, power and other infrastructure, illiteracy, wage labour and low amounts of agricultural assets were associated with poverty. In Kenya, Krishna et al. (2004) found that; poor health, family size and ownership of smaller land holdings were associated with descents into poverty whilst access to urban areas for income diversification and the type of non-farm employment were associated with ascents out of poverty. In Uganda, Lawson et al. (2006) found that; chronic poverty was associated with household composition, average number of cattle owned and average number of average land area owned. Whilst in India, Krishna (2006) found that; poor health, private debt and drought were associated with descents into poverty and diversification of income and improvement of land were associated with ascents out of poverty. In many cases at least some form type of environmental factor are found to be correlated with poverty. These factors could potentially be identified from satellite data. For example, the ownership of small land holdings and increased non-farm jobs due to access to urban areas (Krishna et al. 2004). Estimating the size of agricultural fields (small fields are often an indication of small holder agriculture and a lack of mechanisation associated with larger fields) or estimating the proximity of urban areas to a rural community using data derived from satellite sensors could yield similar results to those found in the above studies.

Associations between poverty and the environment have been studied before in different guises. There is considerable literature on the associations between demography and land use/land cover change (LUCC) which is especially concerned with exploring the colonisation of primary forest into agricultural land in the sub-tropical regions of the World's developing nations (Aldrich et al. 2006; Carr 2004; Cassetti and Gauthier 1977; Rudel et al. 2002). This research area has led to the formulation of several concepts that describe how the environment and demographic conditions are linked. For example, the vicious circle model and the life cycle model (Aldrich et al. 2006; de Sherbinin et al. 2008). Both models have criticised for over-simplifying the relationships between poverty and environment (Gray and Moseley 2005). Often the methods miss out certain important areas in favour of concentrating on one aspect. For example, a study exploring the effects of demographic conditions on land cover change may focus almost entirely on the effects of the population on the natural and physical environments whilst neglecting the reciprocal effect of the environment on the population. Furthermore, it may also neglect to mention the plethora of confounding and mediating effects that can have important impacts on the outcomes being explored (for example, political, economic and institutional factors).

A further area of research that has concerned itself with the links between poverty and the environment is the sustainable livelihoods approach which has been devised as a way of assessing poverty in developing countries without focussing solely on financial capital resources of households (de Sherbinin et al. 2008). The sustainable livelihoods approach states that all households have a range of assets that can typically be divided into five categories; human, natural, physical, social and financial. Each household will have access to a different level of each of these assets. The sustainable livelihoods approach states that households can be financially poor, but still be relatively wealthy due to having access to other forms of capital such as natural resources (for example, woodland and fisheries) and physical capital (for example, ownership of agricultural land). The livelihoods approach demonstrates how socioeconomic conditions and environmental factors can interact to determine a household's well-being and their abilities to deal with changing conditions. These categories indicate the links between environment and poverty and how the two can combine to help determine a households overall wealth.

Poverty and the environment are closely linked but finding and assigning causality between the two is difficult (de Sherbinin et al. 2007) as the associations between rural socioeconomic conditions and environmental factors are complex (Section 2.1.7 discusses some of the uncertainties surrounding these complexities in more detail). This literature review and the subsequent analysis does not attempt to assign causality to the associations, rather it attempted to review the literature that has found links between environment and poverty. This research is therefore, not concerned with the in-depth exploration of links and possible reasons for them. Rather, it focuses on the generic associations between poverty and poor socioeconomic conditions and environmental factors. The review aims to develop further the idea that the presence or absence of particular environmental factors in a rural area may be associated with changing poverty levels of a community. It is acknowledged that this is a large generalisation and oversimplification of the systems at play in the environment-poverty/poverty-environment framework. However, if associations can be found between poverty and the environment using remotely sensed data to identify key environmental metrics it may be possible to use this data to monitor rural areas for potential changes in socioeconomic conditions. It could then be used as a tool to identify areas where further more in-depth research can be conducted and on the sustainable livelihood approach, vicious circle model, and be used for targeting development assistance.

The following sections discuss links and associations between demography, poverty and environment, whereby remotely sensed satellite sensor data could have the potential to be used for identifying environmental factors to monitor socioeconomic conditions in rural communities of developing countries.

2.1.1 Ecological Marginalisation

The poor are often found to be located in areas with unfavourable cultivable land (Sen 2003) due to; (i) environmental changes such as land degradation or loss; (ii) population pressure, and; (iii) a skewed distribution of land ownership that means the majority of the land is owned by absentee land lords (Alam 2003). In South Asia, large scale river bank erosion and subsequent deposition creates river islands or Chars (Sarker et al. 2003). These islands are often inhabited by people who have lost cultivable land through erosion and, due to large population densities and subsequent pressures on the land, are either (i) unable to find suitable cultivable land to purchase, or (ii) unable to afford to purchase land on the mainland. Large numbers of people classed as extremely poor are Chars dwellers in this region of the World (Brocklesby and Hobley 2003). This could be because, although the seasonal flooding on the islands bring highly fertile soils and large yields (Lahiri-Dutt and Samanta 2007; Lein 2008) the sudden loss of land due to larger than normal flooding or more usually bank erosion can result in communities having to prioritize short-term survival over longer term development (Conroy et al. 2010). However, the availability of relatively cheap and highly productive land means that communities choose to live in these marginal lands. The majority would prefer to live on the mainland but could not afford to as for example, in Bangladesh Lein (2008) found that agricultural land on the mainland was up to 10 times more expensive than the equivalent land on the Chars. If unfavourable cultivable land that is susceptible to sudden shocks such as erosion could be identified in remotely sensed data this could be used to find associations with poor socioeconomic conditions.

2.1.2 Forests and woodland

In developing countries the poor are often disproportionately located in forested areas (Sunderlin and Huynh 2005; World Bank 2003). The World Bank estimated that over 1.6 billion people depend on forests for subsistence or economic earnings, with approximately 60 million entirely dependent on forests and 350 million partly dependent (World Bank 2004). This disproportionate pattern can is associated with two main factors:

- Communities with limited alternative land options and economic opportunities choosing to inhabit areas of unclaimed land that can be converted relatively cheaply into agricultural land (Sunderlin et al. 2005), and;
- Indigenous communities remaining in traditional areas (Sunderlin et al. 2007).

One of the largest economic uses of forests in developing countries is timber extraction. However, the majority of timber extraction is conducted by large corporations (PROFOR, 2008; Sunderlin et al. 2004) and can often have very little impact on the development of rural communities (Angelsen and Wunder 2003). Therefore, assessing the true economic impact of forests on rural incomes can be difficult as the majority of rural communities use them as a source of non-timber forest products (NTFP) (Angelsen and Wunder 2003; Dubois 2002).

NTFP often do not alleviate poverty or substantially increase an individual's socioeconomic status (Sunderlin et al. 2007). This is because NTFP are used for domestic subsistence consumption and to supplement incomes during poor agricultural harvests (Sunderlin et al. 2004). Thus, there is "...overwhelming evidence that the poorest segments of the societies around the world are the populations principally engaged in NTFP [non-timber forest product] extraction" (Neumann and Hirsch 2000 p.35). Several studies have found that the collection of NTFP was often on a small-scale, but could make large contributions to household incomes. For example, NTFP have been found to contribute:

- 35% of total household incomes in Zimbabwe (Cavendish 1999);
- 33% of household incomes in Tanzania (Monela et al. 2001);
- Between 17 and 45% of earnings in Bolivia and Honduras (Godoy et al. 2002);
- 22% of total income of rural households across 17 developing countries (Vedeld et al. 2007);
- 39% of total household incomes in south-west Ethiopia (Mamo et al. 2007);
- 27% of total household income in northern Ethiopia (Babulo et al. 2009).

The benefits from forest products can vary spatially depending on local conditions. For example, in northern Ethiopia forest products contributed 35% of the total household income in the poorest quartile of households and decreased to 23% in the wealthiest quartile (Babulo et al. 2009). Similar results were found in south-west Ethiopia where the poorest quintile of households gained 59% of total income from forest products compared to just 34% in the wealthiest quintile (Mamo et al. 2007). It has also been claimed that increased local variations in access to roads and markets can increase the likely benefits from forest products (Angelsen and Wunder 2003; Babulo et al. 2009; Godoy et al. 2002) (roads and markets are described in more detail in Section 2.1.6).

Overall, the presence of open access forest resources in rural areas may help communities maintain socioeconomic standards and provide emergency resources during crop failures. Therefore, areas with a mixture of woodland and other productive land cover types (such as agriculture) may be expected to have positive associations with socioeconomic conditions. However, communities located inside dense forest areas are more likely to have poor socioeconomic conditions. Thus, if woodland and other productive land types can be identified in satellite sensor imagery associations may be found with socioeconomic data that vary spatially.

2.1.3 Water Resources

In developing countries access to irrigation can have a substantial impact on reducing rural poverty (Narayanamoorthy 2007)[3]. The presence of irrigation infrastructure in rural communities has been found to be correlated with lower levels of poverty (Hussain 2007). This can be attributed to irrigation resulting in increased agricultural yields (FAO 1996). An example in China found that irrigated plots had yields that were 177%, 71% and 16% higher than non-irrigated plots for cotton, wheat and maize respectively (Huang et al. 2006). Irrigation has been associated with increased yields through the increased ability of cultivators to grow crops during the dry season which is often not possible using rain fed techniques (Angood et al. 2003; Hussain and Hanjra 2004; Smith 2004). Angood et al. (2003) reported that multiple cropping patterns can lead to surplus crops being sold resulting in increased agricultural incomes. The wider rural community can also benefit from local irrigation as increased productivity can often lead to increased farm and non-farm employment (Smith 2004). Further benefits include increased wages, increased non-farm jobs, decreased food prices, decreased rural out-migration (Bhatterai et al. 2002; IPTRID 1999) and decreased temporary poverty (Hussain et al. 2002).

The impact that access to irrigation can have on rural poverty in developing countries can vary spatially. The impact can be substantial when integrated with investments in other resources such as access to markets and education (Hussain 2005). For example, the construction of main roads in Nepal was thought to enhance the impacts of irrigation (Angood et al. 2003). However, Van de Walle (2003) reported that poor education could limit the benefits of irrigation.

---

[3] Detailed reviews of the impacts of irrigation can be found in Smith (2004) and Lipton (2007)

Therefore, if irrigation or irrigated crops can be identified in satellite sensor imagery associations may be found with socioeconomic variables. The associations would however be expected to vary spatially. Thus, analysis would have to account for other environmental conditions such as road availability (Section 2.1.6).

2.1.4 Agriculture

In developing countries the agricultural sector can often have the most significant impact on rural poverty and development than any other economic sector (Cervantes-Godoy and Dewbre 2010). Of particular importance to rural economic development is growth in the agricultural sector and increased agricultural yields (Datt and Ravallion 1998; Ligon and Sadoulet 2008; Ravallion and Chen 2007; Thorbecke and Jung 1996). This positive impact centres around increased rural employment, growth in the rural economy and decreased rural food prices (de Janvry and Sadoulet 2010; Irz et al. 2001). Thus, a lack of agricultural development and growth can often have a detrimental effect on poverty and rural development (Christiaensen and Demery 2007; Christiaensen et al. 2006). Furthermore, in developing countries, the growth in agricultural production, and rural growth in particular, has been found to reduce poverty in both rural and urban areas whereas urban growth has little impact on rural poverty (Ravallion and Datt 1996).

If agricultural development over time or levels of agricultural yield across space could be identified in satellite sensor imagery, associations may be found with socioeconomic data. However, other aspects of development can affect the impact that agricultural development and growth can have on poverty reduction. For example, changes in government policy appeared to have a larger impact on economic growth and poverty reduction than agriculture in Brazil (Ferreira et al. 2010) and China (Ravallion and Chen 2007). Other issues such as land policy and equitable access to agricultural land can also have an impact on the levels of development associated with agricultural growth (DFID 2002a: Hussain 2007; Lipton 2007). Infrastructure, such as the quality of rural roads, may also have an association with agricultural development. Research indicates that the level of agricultural growth and its effect on rural development would be lower without infrastructure (Mellor 1999; Serneels and Lambin 2001). However, this is a controversial issue and is discussed in more detail in section 2.1.6 and 2.1.7.

2.1.5 Biodiversity

Maintaining biodiversity is often seen as an important factor for rural development and poverty alleviation (Campbell et al. 2002; FAO 1996; Hussain and Hanjra 2004; IFAD 2001; IFAD 2009; Ligon and Sadoulet 2008). Biodiversity preservation is a key contributing factor in the poverty reduction strategies of international bodies such as the UN and FAO (Tallis et al. 2008). The importance of biodiversity for rural development was highlighted by the United Nations Environment Programme (UNEP):

> "*Maintenance of a heterogeneous local environment provides the widest possible range of ecosystem services reduces the exposure of local people to risk and lessens their dependence on the vagaries of global markets or on development assistance*." (2007 p.2)

The rural poor are often dependent on natural resources directly. Therefore, maintaining biodiversity and balancing this with agricultural development can be important for the rural poor to ensure future food security. In Thailand it was suggested that conserving and protecting areas had helped to reduce poverty in the local areas through increased tourism (Sims 2010). However, protecting biodiversity can have negative impacts on rural development as it can limit the local uses of land that was once used as a common resource (Adams et al. 2004; IUCN 2008). A review of 32 World Bank biodiversity projects carried out between 1993 and 2007 by Tallis et al. (2008) found that just five (16%) of the projects had substantial positive impacts on poverty alleviation. If remotely sensed satellite sensor imagery can be used to estimate the biodiversity of a region or identify the amount of protected land in an area this could be used to explore associations with poverty and economic development.

2.1.6 Roads and Accessibility

Community remoteness can result in decreased access to important facilities such as non-agricultural employment, health, education and markets (ADB 2001; Blaikie et al. 2002; Hanmer et al. 2000; IFAD 2001; Khandker et al. 2006; Porter 2002). This can often result in larger concentrations of rural poverty in remote or marginalised areas of developing countries (IFAD 2002; Porter 2002; Sen 2003) and less remote communities experiencing more rapid poverty reduction in developing countries (Bird and Shepherd 2003).

In rural areas of developing countries research has indicated that roads are the most common method of accessing markets, and other important facilities such as health and education. Therefore, they have been found to be associated with decreased levels of poverty and increased agricultural productivity and household incomes (IFAD 2002; Lokshin and Yemstov 2005; Platteau 1996; Windle and Cramb 1997). Research has found that agricultural inputs, such as fertiliser and pesticides, which can increase crop yields, are used more in areas with easier access to markets (Von Oppen et al. 1997). This is often because easier access to roads effectively reduces the cost of commercial inputs such as fertiliser, pesticides and higher yielding seed varieties which can then result in increased uptake of these inputs and can lead to increased farm employment (Fan and Hazell 2001; IFAD 1999).

Diversification of rural incomes from reliance on agriculture to non-farm incomes can contribute to a reduction in rural poverty (Lanjouw and Lanjouw 2001). Rural roads can play a vital role in the diversification of incomes because, as well as benefitting the agricultural sector rural roads can provide access to alternative forms of income. However, non-farm incomes will not reduce or remove poverty for all participants (van de Walle and Cratty 2004) and they can contribute to increased inequality (Canagarajah et al. 2001). This may be partly explained by a study by Lanjouw (2001) in El Salvador which found that those with diverse incomes from farm and non-farm sources were more likely to be poor compared to those engaged only in non-farm activities. This emphasises the complexities in exploring development and poverty reduction but also that some actions such as income diversification are driven by necessity to survive rather than as a way to increase overall living standards (Lanjouw and Lanjouw 2001).

The direct and indirect benefits that roads can bring to rural development have led several international non-governmental organisations (NGOs) to give high importance to rural road investment schemes. For example, the World Bank invested $9.4 billion in international transport initiatives in 2010 and over the last decade supported over 200 projects worth over $34 billion or 21% of the Banks portfolio (World Bank 2011a). Further, a large proportion of the development loans handed out by the Asian Development Bank between 1986 and 2001 supported road development (ADB 2001).

Not all rural road schemes result in increased prosperity for rural communities. Often only portions of the rural population are able to benefit from the schemes (Ellis 1998; Wilson 2004) as the poorer members of rural communities lack the financial and human capital to diversify farm and non-farm incomes. Rigg (2002) summarised the complexities in assessing the impact of rural road schemes on poverty reduction; "…there is a range of cultural, social and economic reasons – which often coalesce – why roads offer unequal opportunities to different groups" (p 630). Furthermore, Escobal (2001) found that access to roads was one of a number of issues that could explain increased farm and non-farm incomes in rural Peru. However, the presence of rural road schemes has been found to contribute significantly to rural development. Therefore, if rural roads can be identified and access to markets estimated using satellite sensor imagery this information could be used to explore the associations with socioeconomic variables. However, it is important to consider that the poorest communities do not always benefit the most from rural road schemes as transportation costs can remain high (Jacoby 2000). Separating the impact that different interventions have on the community can be difficult and therefore the importance that each aspect has with poverty reductions and rural development is uncertain (Khandker et al. 2006; Leinbach 1995; Van de Walle 2002).

2.1.7 Uncertainties in the associations between socioeconomic variables and the environment

> "*Although there is widespread cognition that important relations exist between population, development and environment, there is little agreement about the nature and magnitude of the links*" (UNFPA 2001 p.11).

This statement highlights the links between human development and environmental factors. However, these links can be complex and it can be difficult to assign causality to the two issues. Taking the comments by Rigg (2002) in section 2.1.6 further, Figure 2.1 summarises some of the main confounding and mediating effects in the associations between socioeconomic conditions and environmental factors. The figure demonstrates that population and environmental variables can be directly linked (blue line). However, there are several additional factors (central boxes) that can have mediating or confounding effects on these links. Furthermore, there can be additional associations between each of the mediating and confounding effects. For example, as the previous sections have demonstrated economic development is often expected to increase as access to facilities becomes easier (direct links). However, this association cannot be expected to be linear as government policy could result in some communities utilising access to facilities in different ways. For example, government subsidised transport would encourage more people to travel to towns and cities for markets compared to an area with no subsidised transport. Furthermore, government policy in one area may be affected by the cultural conditions as they seek more votes from particular sections of society which will affect local government policy.

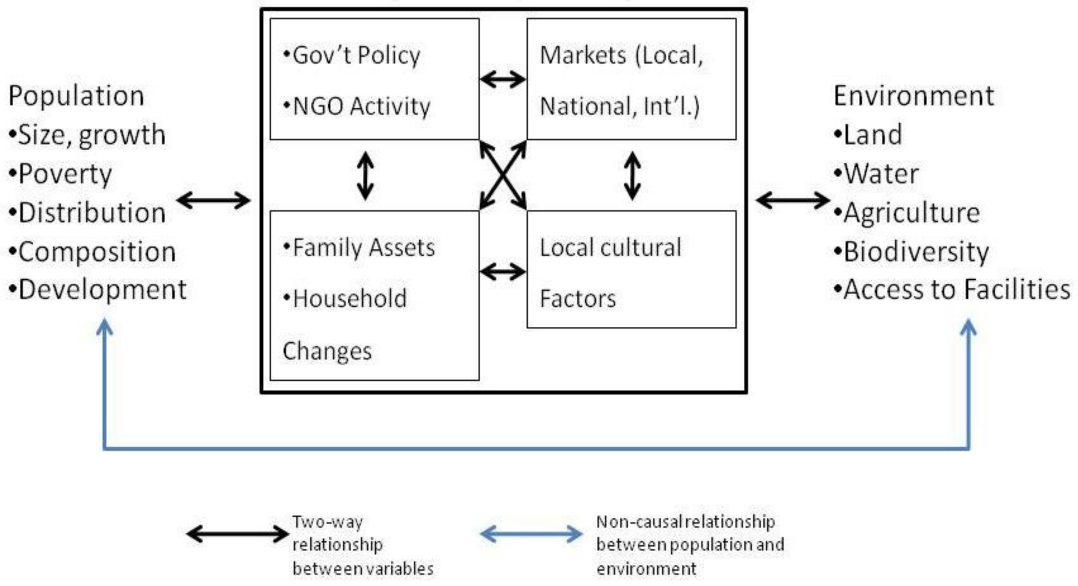

Figure 2.1 complex associations between population and environment[4].

Confounding and mediating factors in the links between poverty and the environment can include the type of local environmental conditions present, amount of assets held by the household (Reardon and Vosti 1995), the country of study (Dasgupta et al. 2001) and government and market failures (Duraiappah 1998). Associations can also change as development progresses (de Sherbinin et al. 2008) and have been found to vary spatially within and across countries (de Sherbinin et al. 2008; Reardon and Vosti 1998; Sen 2003). Furthermore, it has been suggested that integrated investment in several different aspects of the rural economy would create more effective poverty alleviation (Hanjra et al. 2009).

The links between population and environment in relation to tropical deforestation have often concluded that the historical simplification that; poverty drives deforestation through shifting cultivation (Lambin et al. 2001). However, the relationships between population and environment are more complex than this statement suggests (Geist and Lambin 2002). For example, Lambin et al. (2001) found that institutional policies fed into local and national markets which created opportunities for the population to take advantage of. Similarly, Geist and Lambin (2002), Lambin et al. (2003) and de Sherbinin et al. (2007) highlight the complexities that should

[4] Figure adapted from UNFPA, (2001)

be considered when exploring links between poverty and the environment. For example, deforestation can be the result of increased land use pressure due to increased population. However, it can also be linked to market demand, government investment, changes in technology, and subsidies and changes in land policy (Figure 2.2).

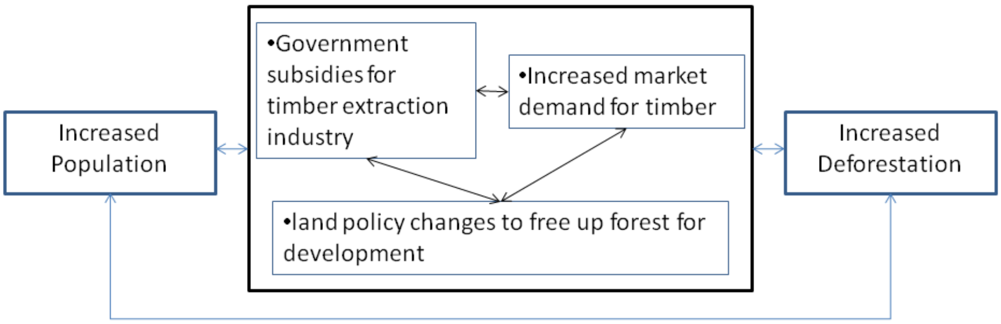

Figure 2.2 Visualisation of the links between population and deforestation with additional mediating affects in the centre.

The scale of analysis can be a very important consideration when exploring the associations between population and environment. The majority of the research discussed in Section 2.1 was focused at the village level. Such small area findings may lead to important macro scale impacts, such as regional government development policies, being missed from the analysis. Therefore, important associations may be missed unless data cover all spatial scales (local, regional and national). This lack of multi-scale information could lead to the identification of associations that may not have resulted in significant results, had other variables been available. Consequently, this is a primary reason why causal associations cannot be attributed to population and environment (Young et al. 2006) as it is unlikely that data at all spatial scales will be available for analysis. Therefore, no attempt will be made to attribute causal associations in this research. Instead this research explores possible associations between socioeconomic conditions and environmental factors to determine if socioeconomic conditions can be predicted using remotely sensed environmental metrics.

2.1.8 The value of remote sensing for exploring associations between socioeconomic variables and environmental factors

Individual environmental resources can provide important benefits to rural populations of developing countries. As population numbers increase and demand for food, water and resources also increases, pressures on environmental resources may lead to resource degradation (Lambin and Meyfroidt 2010). As populations become more dynamic and population change becomes much more rapid it could be argued that the data contained in large scale census and sample surveys will become less useful for policy development and vulnerability assessment.

Earth observation satellite sensors can collect information about land cover and its changes at high temporal and spatial resolution. If the associations between population and the environment highlighted in Section 2.1 can be found using socioeconomic data derived from a census and environmental information derived from Earth observation data it may be possible to provide more up-to-date estimates of socioeconomic conditions. This updated information could be used to identify areas where resources should be targeted.

It is clear that methods to explore the associations between socioeconomic conditions and environmental information will have to account for the complexities that have been highlighted throughout Section 2.1. Methods will have to explore:

- How to link census data and remotely sensed imagery;
- The scale at which the analysis is conducted, and;
- How to account for and explore spatially varying associations between poverty and the environment.

2.2 Methods to link census derived socioeconomic data and remotely sensed environmental factors

To explore the associations between socioeconomic variables and remotely sensed environmental factors the spatial boundaries in which the population and environment interact have to be defined (Rindfuss et al. 2003a). It is common for socioeconomic data and environmental data to be stored in different formats. Socioeconomic data are most often represented in vector format reflecting the scales of data collection. These can include, discrete points for individuals, households, or settlement centroids or polygons representing village boundaries, census enumeration districts and counties. Conversely, remotely sensed environmental data are stored in raster format representing continuous coverage within a defined area. Socioeconomic vector data can be linked with remotely sensed environmental data to provide information about the local environment using a geographic information system (GIS). A GIS can be used to extract local environmental information and spatial analysis performed to relate environmental information to socioeconomic data.

Several methods exist to link census derived socioeconomic data with environmental factors derived from remotely sensed satellite sensor data. Methods used depend on the purpose of the study, the spatial scale of the socioeconomic data, the spatial scale (resolution) of the satellite sensor data, and, the spatial extent of the available socioeconomic data (Rindfuss et al. 2003b).

2.2.1 Linking individual households with individual land parcels

The most appropriate scale of linking social and environmental data is often at the household – level because many socioeconomic data, such as the census and sample surveys, are collected at this scale. For household-level socioeconomic data environmental factors that could be associated with individual households need to be identified at the same scale. When households are situated on land that is cultivated by the household and fine spatial resolution satellite sensor data are available the socioeconomic and environmental data can be linked at the household level (Marquette 1998; McCracken et al. 1999). Figure 2.3 shows that it would be possible to identify the types of land within the boundary of a household. This would enable the amount of productive land, agricultural yield, access to infrastructure and access to road infrastructure within each household to be related to socioeconomic conditions which would provide detailed analysis into the associations between socioeconomic conditions and environmental factors. However, extensive field work would be required to collect global positioning system (GPS) coordinates at the four corners of each land parcel (Walsh et al. 2003).

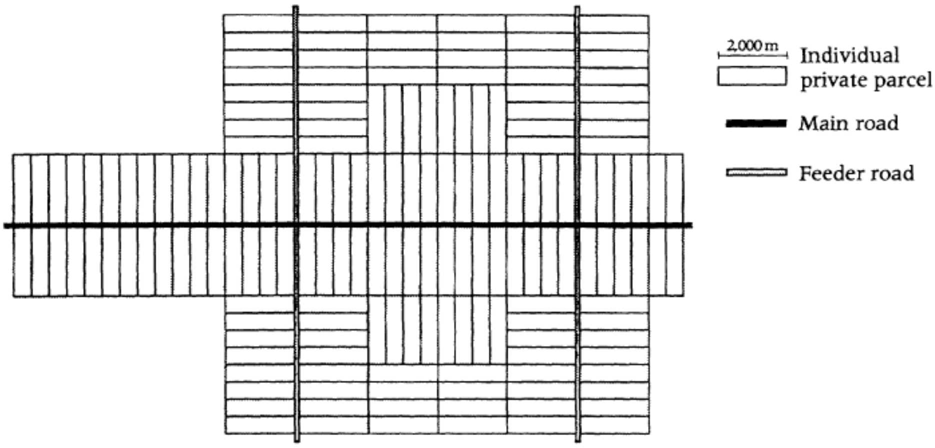

Figure 2.3 Fishbone settlement pattern in South America, source: Evans and Moran (2002 p.176).

A problem for linking socioeconomic and environmental data at the household level is land fragmentation. In areas with rapid population growth land is often divided between children or sold off (Pan et al. 2004). The resulting irregularly shaped land parcels can complicate linking methods (Figure 2.4). However, land registry maps could be used to link individual households with each non-contiguous land parcel that they own (Rindfuss et al. 2003b). These methods would create an accurate portrayal of land ownership. However, they would increase the financial and time costs of field data collection.

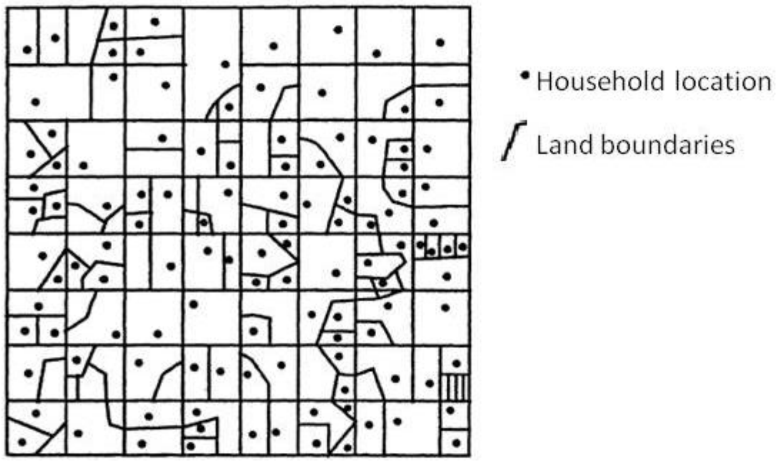

Figure 2.4 Complexities of linking individual households to individual land parcels in North America, source: Evans and Moran (2002 p.179).

Linking socioeconomic data and environmental factors at the household level would reflect the scale at which most land-use decisions are made. It would also most likely be the level at which changes in environmental conditions would have the greatest impact on socioeconomic conditions. However, this method would only be feasible for small scale studies as the amount of field data and the fine spatial resolution of satellite sensor data required would be vast, expensive and time consuming for subsequent analysis.

2.2.2 Linking aggregated socioeconomic data with wider scale landscape data

Although census data are collected at the household level it is not always necessary or appropriate to link socioeconomic and environmental datasets at this level (Lambin 2003) or possible due to aggregation of data to maintain anonymity. When individual socioeconomic data are aggregated alternative linking strategies are required. Radial buffer zones created can provide an approximation of village boundaries which can then be spatially related to the environmental data that is most likely to be associated with socioeconomic conditions (Entwisle et al. 1998).

A study by Behrens et al. (1994) into potential reasons for land use intensification in Venezuela used radial buffer zones to represent the boundaries of villages and to link socioeconomic and environmental data sets. A GPS was used to locate village centres and the edges of the fields furthest from the village centre. The GPS points were used to create polygons in a GIS with different sizes for each village. Subsequent analysis was performed using environmental data that had been extracted within each buffer zone (Behrens et al. 1994). A more generalised method was used in studies in Thailand where socioeconomic data were available at the village level and represented in a GIS by a GPS point for each village centre. Buffer zones with a radius of between 2 and 3 km were used to derive local environmental data (Entwisle et al. 1998; Rindfuss et al. 2003b). Unlike the method employed by Behrens et al (1994) all of the buffer zones created by Entwisle et al. (1998) and Rindfuss et al. (2003b) were the same size and required no field data collection of village boundary information. The method would lead to more generalised results than those expected in Behrens et al. (1994) as it was assumed that the same amount of land was available within each village which is unlikely to be the case. However, this was an efficient way of deriving an estimate for the amount of land around each village and would be more applicable for research on extensive spatial scales.

Radial buffer zones can provide an effective and rapid method of linking socioeconomic data with environmental data. However, buffer zones will inevitably produce omission and commission errors related to ownership and use of land. Figure 2.5 highlights these errors when land owned by village A is outside of the village buffer zone or land owned by village A is inside the buffer zone of village B (Rindfuss et al. 2003b). The land holdings shaded in black in the left hand village belong to the household coloured black and are all within the radial buffer. They therefore would be linked to the correct village in subsequent analysis. However, black shaded land parcels in the right hand village are outside the buffer and therefore will not be considered in subsequent analysis which could cause bias in the data. Red shaded land parcels, which belong to the red household, are distributed across the study area and, in one case, a field is inside a different village buffer zone. These errors could affect the results of any subsequent analysis as the socioeconomic data will not match the relevant environmental data. Further, associations between socioeconomic conditions and environmental factors can extend far beyond the local community. Thus, exogenous environmental factors that have potentially important associations with socioeconomic conditions may be missed when using buffer zones (Evans and Moran 2002).

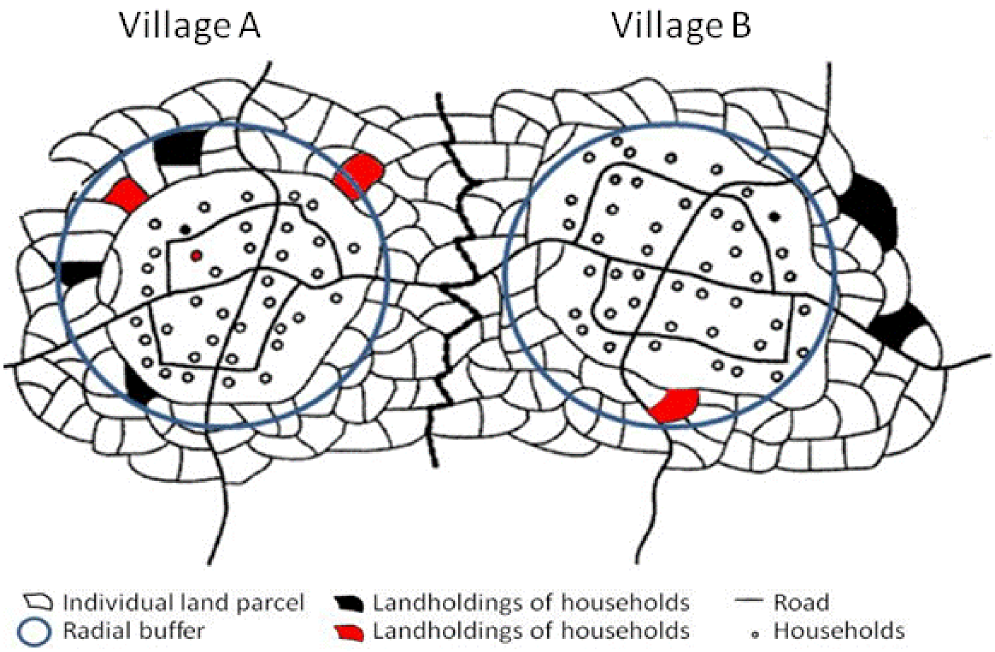

Figure 2.5 Land misrepresentations associated with radial buffers which can induce errors (adapted from Evans and Moran, 2002 p.182).

The size of radial buffer zone sizes could be increased to capture more environmental information and reduce the omission errors. However, this could lead to increased commission errors and replication of environmental data resulting from overlapping village extents. An alternative method could be used where the size of each buffer zone was weighted to represent a more appropriate size of different villages. This could be done using socioeconomic data such as the population size of individual villages or data indicating the ownership and location of every land parcel. However, this would not be feasible for research on large spatial areas. Further, radial buffers could result in large amounts of environmental data not being assigned to any socioeconomic data level and thus not being included in subsequent analysis.

When exploring the dynamics of land use change across an extensive area of Vietnam, Muller and Zeller (2002) used Thiessen or Voronoi polygons (Longley et al. 2011) to integrate socioeconomic and environmental data. Unlike radial buffer zones a Thiessen polygon

"…contains only one input point [village], and any location within a polygon is closer to its associated point than to the point of any other polygon" (ESRI 2009a). Thus, all environmental data across the study area would be assigned to the closest socioeconomic data point. This would potentially reduce omission and commission errors associated with radial buffers in rural areas of developing countries as all environmental data would be included in the analysis once. In the absence of village boundary information Thiessen polygons may be more appropriate for analysis in rural areas of developing countries as, unlike developed countries, large amounts of land are used even if they have protected status due to informal land management practices and policies. However, the use of Thiessen polygons could potentially generalise the associations between socioeconomic and environmental data too far by only allowing environmental data to be assigned to one socioeconomic data point.

2.2.3 Summary of linking strategies

The most appropriate method of linking socioeconomic and environmental data will depend on the level of spatial aggregation of the socioeconomic data and the purpose of the subsequent analysis. Census data are collected at the household level and most commonly aggregated to census enumeration districts and thus analysis is limited to this spatial scale (Lo 1997). However, when socioeconomic data are available at finer resolutions such as the household or village level different methods can be employed. Linking at the household level can ensure the most detailed information is retained for subsequent analysis but requires large amounts of data and would most likely be infeasible for many studies exploring links across extensive spatial areas.

Radial buffer zones appear to offer a rapid and effective way of linking aggregated socioeconomic data and environmental datasets. Methods for limiting the omission and commission errors could include weighting the size of the buffer zone. However, the aggregated data resulting from buffer zone analysis will not be able to account for all of the heterogeneity within the village (Lambin 2003) and this would have to be considered when interpreting the results of subsequent analysis. Overall, the linking strategy would be primarily determined by the spatial scale of the socioeconomic data and the spatial extent of the study region. A consideration when exploring the associations between socioeconomic and environmental data is that the linking strategy could have an important influence on the final results. Therefore, any study using generalised linking strategies should seek to explore the impact that alternatively sized buffer zones would have on the final results of the analysis.

2.3 Statistical methods to identify potential associations between socioeconomic variables and environmental factors.

Statistical and spatial statistical models can be used to explore and identify associations between poverty or socioeconomic variables and remotely sensed environmental metrics. Statistical methods include multivariate approaches that explore the significance and magnitude of associations between an observation variable and a set of explanatory variables which in the case of this research would be a socioeconomic indicator and environmental factors respectively (Crews and Walsh 2009).

Models to explore the significance and magnitude of associations between different variables include correlation analysis and multiple regression analysis (Chapter 6). Correlation analysis compares the variation in two variables and the resulting correlation coefficient provides an estimate of the association between the two variables (Reimann et al. 2009). Regression analysis advances this method by providing an estimate of the power of the explanatory variables to predict the observed variable (Dalgaard 2002; Field 2009; StatSoft Inc. 2007). Regression analysis aims to explain how the variation in the dependent variable depends on the variation in independent variable(s). The simplest forms of regression analysis (linear regression) achieve this by finding the straight line model that passes as close to the majority of the data points as possible (Field 2009). The correlation coefficient from this analysis can be squared (r^2) to create a coefficient of determination, which can be used to show how much variability in the observed variable is accounted for by the independent variable(s) (Chapter 6 provides a more detailed description of non-spatial regression modelling; Chapter 7 and 8 provide more detail on spatial regression modelling).

2.3.1 Spatial Analysis

Associations between poverty and environmental factors have been found to vary across space (Section 2.1). Traditional regression analysis assumes that the associations between variables being regressed do not vary across the study area. If associations do vary across the study site simple regression methods can be inappropriate for the analysis as they only generate an average model and do not account for local effects. Therefore, regression models will be extended in this study to explore the effects that spatial dependency (Chapter 7) and spatial non-stationarity (Chapter 8) have on the model results.

2.4 Summary

Chapter 1 identified some of the limitations of census data for use in targeting government resources and academic research in poverty alleviation, economic development and social vulnerability to natural hazards. Also identified was the potential for high resolution satellite sensor imagery to be used to update the census between enumeration periods.

This chapter has highlighted that there are known and measurable associations between socioeconomic variables and environmental factors in the literature. These associations are however, mediated and confounded by a range of complexities and uncertainties often related to the causality of associations and spatially heterogeneous patterns across space. However, these associations and complexities can be taken into consideration using spatial analysis techniques.

The aim of this research was to explore and statistically quantifying the spatial associations between census based socioeconomic variables and environmental factors obtained from fine resolution satellite sensor imagery. This chapter has introduced the main theoretical basis for the thesis and highlighted some of the methods that will be further explored a detailed description and review of the methods is presented in each of the analysis chapters.

2.5 Research Objectives

This thesis focuses on exploring associations between census based socioeconomic variables and environmental factors derived from remotely sensed satellite sensor imagery. From the literature reviewed and the contextualization discussed in the introduction the following research objectives have been identified:

Objective 1: Explore if associations in the literature translate to a case study location
Identify a suitable case study location to explore the local associations between socioeconomic conditions and environmental factors. Use this case study location to determine if associations highlighted in the literature are prevalent in this locality and identify which, if any, associations between socioeconomic conditions and environmental factors vary spatially. Finally, use field observation to consider how remotely sensed satellite sensor data could be used to generate proxies for the environmental factors associated with socioeconomic conditions.

Objective 2: Develop targeted remotely sensed environmental variables
Few of the links between socioeconomic and environmental data highlighted in Section 2.1 are directly observable from remotely sensed satellite sensor data. Therefore, this objective addresses the development of a targeted land cover classification scheme that generates environmental variables; variables which have potential associations with local socioeconomic conditions as identified from field observations and the reviewed literature.

Objective 3: Explore the associations between socioeconomic and environmental variables using global regression techniques
Explore and quantify the associations between socioeconomic conditions and environmental metrics using a global non-spatial regression technique. The results of this analysis will be used as a benchmark to assess the general associations across the case study area prior to conducting more detailed spatial analysis.

Objective 4: Explore how spatial dependence in model errors can affect the estimated associations

The global statistical models used in Objective 3 make assumptions about the data and associations between different datasets that are often violated when using socioeconomic and environmental data. These violations can result in spatial autocorrelation (SAC) which can, in turn, result in unreliable associations between socioeconomic conditions and environmental factors. This objective will investigate if global regression results display SAC and determine how accounting for SAC impacts the prediction of socioeconomic conditions using environmental factors.

Objective 5: Explore if associations between socioeconomic data and environmental factors vary spatially

The literature has highlighted that associations between socioeconomic and environmental variables can vary spatially within a region of interest. A spatial statistical method will be used to explore if spatially varying associations exist across the study region. In addition, it will be explored how accounting for the spatial non-stationarity of the model can affect the prediction of socioeconomic conditions from remotely sensed environmental metrics.

Objective 6: Explore methods of estimating community access to important facilities

The literature has highlighted the importance of road infrastructure and access to facilities such as markets, education and health for rural development. An objective will be to explore the most appropriate methods for representing access to important facilities using only satellite sensor derived information. Considerations need to be made for models accounting for road densities, estimates of travel effort from village to urban areas using cost-surfaces and estimates of distances from villages to urban areas using straight line distance measures.

Objective 7: Explore the most appropriate models for linking census and environmental data

As highlighted from the reviewed literature, one of the most important aspects of linking socioeconomic and environmental data is the method used to link the datasets which is often determined by the data available. Therefore, the most appropriate models for transforming environmental variables will be explored such that, the association with socioeconomic variables will be maximised. Such models to explore include fixed size radial buffer zones, weighted size buffer zones and Thiessen polygons. Models will identify the environmental variables within the vicinity of the social element (for example, administrative boundaries, villages, individual households) to link to census variables.

These research objectives form the basis for the subsequent research-based chapters of the thesis. Objectives are addressed in Chapters 5 to 8 and overall linkages between objectives are discussed in Chapter 9. A final conclusion in Chapter 9 summarises the key findings of the research objectives and provides a consensus as to the overall research aim of how useful environmental factors acquired from fine spatial resolution remotely sensed imagery are for estimating socioeconomic data.