University of Turku, Department of Geography and Geology

Markus Kukkonen

FOREST COVER AND ITS CHANGE IN UNGUJA ISLAND, ZANZIBAR

Geography thesis

Keywords: Forest change, deforestation, remote sensing, change detection, GIS, regression analysis, Zanzibar, Tanzania

Turku 2013

UNIVERSITY OF TURKU
Faculty of Mathematics and Natural Sciences
Department of Geography and Geology

KUKKONEN, MARKUS: Forest cover and its change in Unguja Island, Zanzibar

Master's thesis, 166 p., 6 appendix pages.
40 ECTS
Geography
January 2013

Tropical forests are sources of many ecosystem services, but these forests are vanishing rapidly. The situation is severe in Sub-Saharan Africa and especially in Tanzania. The causes of change are multidimensional and strongly interdependent, and only understanding them comprehensively helps to change the ongoing unsustainable trends of forest decline. Ongoing forest changes, their spatiality and connection to humans and environment can be studied with the methods of Land Change Science. The knowledge produced with these methods helps to make arguments about the actors, actions and causes that are behind the forest decline.

In this study of Unguja Island in Zanzibar the focus is in the current forest cover and its changes between 1996 and 2009. The cover and changes are measured with often used remote sensing methods of automated land cover classification and post-classification comparison from medium resolution satellite images. Kernel Density Estimation is used to determine the clusters of change, sub-area –analysis provides information about the differences between regions, while distance and regression analyses connect changes to environmental factors. These analyses do not only explain the happened changes, but also allow building quantitative and spatial future scenarios. Similar study has not been made for Unguja and therefore it provides new information, which is beneficial for the whole society.

The results show that 572 km^2 of Unguja is still forested, but 0,82–1,19% of these forests are disappearing annually. Besides deforestation also vertical degradation and spatial changes are significant problems. Deforestation is most severe in the communal indigenous forests, but also agroforests are decreasing. Spatially deforestation concentrates to the areas close to the coastline, population and Zanzibar Town. Biophysical factors on the other hand do not seem to influence the ongoing deforestation process. If the current trend continues there should be approximately 485 km^2 of forests remaining in 2025. Solutions to these deforestation problems should be looked from sustainable land use management, surveying and protection of the forests in risk areas and spatially targeted self-sustainable tree planting schemes.

Keywords: Forest change, deforestation, remote sensing, change detection, GIS, regression analysis, Zanzibar, Tanzania

TURUN YLIOPISTO
Matemaattis-luonnontieteellinen tiedekunta
Maantieteen ja geologian laitos

KUKKONEN, MARKUS: Metsänpeite ja sen muutos Ungujan saarella, Sansibarilla

Pro gradu -tutkielma, 166 s., 6 liites.
40 op
Maantiede
Tammikuu 2013

Trooppiset metsät tuottavat monia keskeisiä ekosysteemipalveluita, mutta näiden metsien määrä vähenee hälyttävällä tahdilla. Metsien väheneminen on nopeaa Saharan eteläpuolisessa Afrikassa ja erityisesti Tansaniassa. Syyt muutoksiin ovat monitahoisia ja vahvasti keskinäisriippuvaisia, mutta ainoastaan ymmärtämällä niitä kokonaisvaltaisesti voidaan muuttaa nykyisiä kestämättömiä kehityskulkuja. Metsissä tapahtuvia muutoksia, muutosten spatiaalista rakennetta ja suhdetta ympäristöön ja ihmisiin voidaan tutkia maanmuutostieteen metodein. Näillä työkaluilla tuotettu tieto auttaa ymmärtämään ketkä ovat niitä tekijöitä ja mitkä ovat niitä syitä jotka muokkaavat maailman metsiä.

Tutkimuksessa keskitytään Sansibarilla sijaitsevan Ungujan saaren metsien nykyiseen määrään ja siinä vuosien 1996 ja 2009 välillä tapahtuneisiin muutoksiin. Metsien määrää ja muutoksia mitataan kaukokartoituksesta tutuilla automaattisen maapeitteen luokittelun ja muutosanalyysin menetelmillä. Tiheysfunktion ydinestimaattia käytetään muutosten spatiaalisten klustereiden kartoittamiseen, osa-alue – analyysillä tuodaan esiin alueiden välisiä eroja kun taas etäisyys- ja regressionanalyysit yhdistävät tapahtuneita muutoksia ympäristötekijöihin. Näiden analyysien avulla ei ainoastaan selitetä tapahtuneita muutoksia vaan pyritään myös ennustamaan miten paljon ja missä metsäkatoa tapahtuu tulevaisuudessa. Samankaltaista tutkimusta ei ole aiemmin tehty koko Ungujan alueelta, joten se sisältää paljon uutta tietoa mistä on hyötyä koko yhteisölle.

Tulokset näyttävät, että noin 570 km^2 Ungujasta on vielä tänäänkin metsien peitossa, mutta 0,82–1,19% näistä metsistä katoaa vuosittain. Horisontaalisen metsäkadon lisäksi myös vertikaalinen degradaatio ja spatiaaliset muutokset ovat merkittäviä ongelmia. Metsäkato on intensiivisimmillään yhteisömailla olevissa alkuperäismetsissä, mutta myös peltometsäalueet vähenevät. Metsäkadon spatiaalista rakennetta selittää rannikoiden, asutuksen ja Zanzibar Townin läheisyys, kun taas biofyysiset tekijät eivät vaikuta prosessiin. Jos kehityskulku jatkuu nykyisellään, vuonna 2025 metsät peittävät enää vain noin 485 neliökilometriä saaresta. Ratkaisuja metsäkadon ongelmiin tulisi hakea kestävämmästä maankäytönsuunnittelusta, riskialueilla sijaitsevien metsien kartoittamisesta ja suojelusta sekä alueellisesti kohdistetuista omavaraisista puunistutushankkeista.

Avainsanat: Metsien muutos, metsäkato, kaukokartoitus, muutosanalyysi, paikkatieto, regressioanalyysi, Sansibar, Tansania

Table of contents:

1. Introduction

Tropical forests are known to sustain biodiversity and hydrological cycle, improve air, water and soil quality, slow down climate change, reduce flooding and erosion, uphold traditional cultures and provide livelihoods to communities and individuals (Skole & Tucker 1993; Rudel & Roper 1997; Laurance 1999; Houghton et al. 2000; Maass et al. 2005; UNEP 2007; FAO 2010: 35–37; Fagerholm et al. 2012). Nonetheless 5,2 million hectares of forests, area larger than Estonia, is lost each year. Even though the public discussion often focuses on South America, the forests of Africa are almost as threatened (FAO 2010: 15–34). Tanzania is one of the deforestation hot spots of the continent and as part of Tanzania, Zanzibar is no exception (RGZ 2004; DCCFF 2008; FAO 2010: 21).

Deforestation and forest changes in general are outcomes of multiple processes. Different actors cause change at local level because of wood extraction, agricultural and infrastructure expansion, but the amount of deforestation is directed by more underlying causes in the society and the environmental realm makes the boundaries for how much and where deforestation takes place (Rudel & Roper 1997; Kaimowitz & Angelsen 1998: 90–98; Angelsen 1999; Geist & Lambin 2001: 6–8). The outcomes of deforestation can be measured with remote sensing and GIS techniques, but understanding the processes and causes behind these measurements requires approaches incorporated from multiple disciplines (Nagendra et al. 2004; Rindfuss et al. 2004; Lambin et al. 2006: 7; Turner et al. 2007). These interdisciplinary approaches are referred as Land Change Science (LCS). LCS tries to provide answers to questions how land can change, how people cause change and how the environmental factors distribute the change spatially (Rindfuss et al. 2004; Lambin et al. 2006: 8; Turner et al. 2007).

Changes are usually measured with post-classification comparison methods in many LCS studies (Lu et al. 2004; Pontius et al. 2004). However the focus has shifted away from simply detecting changes to explaining and modeling them. This is often done with change trajectory, distance, multivariate regression and cellular automata analysis (Mertens & Lambin 2000; Veldkamp & Lambin 2001; Verburg et al. 2004; Käyhkö & Skånes 2006). These studies have been made from various areas around the world and they have deepened the understanding about the causes of forest decline, provided knowledge about the sustainability of current forest uses for the stakeholders, allowed discussions about alternative options and provided tools for decision makers to allocate actions (Ludeke 1990; Chomitz & Gray 1996; Mertens & Lambin 1997; Verburg et al. 2002, 2004, 2006; Walker 2004).

This study tries to monitor, model and explain the forest change processes happened in Zanzibar between 1996 and 2009 with the theoretic and methodological framework provided by LCS and deforestation research. The forests of Zanzibar are part of biodiversity rich East African Coastal Forest and they are homes to various endemic species (Burgess & Clarke 2000: 71–73). Also the daily lives of many Zanzibarians depend on forest products as sources of energy and livelihoods (Sitari 2005; Fagerholm & Käyhkö 2009; Fagerholm et al. 2012). However these forests are threatened by intensifying shifting cultivation, unsustainable fuel wood collection, expansion of tourism, urbanization, population growth and unsecure land tenure (RGZ 2004: 6; DCCFF 2008; Käyhkö et al. 2008: 73–74; Mustelin 2008; Käyhkö et al. 2011). The authorities of the archipelago are seriously concerned about the sustainability of the natural resources (DCCFF 2008). Also Finland has collaborated with the Government of Zanzibar since 1980s in the field of forestry and securing the existence of forests have been one of the main goals of this collaboration (Fagerholm 2012: 35).

Estimations about the extent of current forest cover and the rate of its decline have been made, but these have been based on spatially limited field observations or they are badly outdated (RGS 2004; DCCFF 2008). There is a lack of spatially explicit research, which would survey the coverage of current forests, provide relatively accurate rates for deforestation and map the spatial patterns of lost and still existing forests. This research tries to answer to these needs for Unguja, the main island of Zanzibar. The goal is not only to detect the changes at the level of entire Unguja, but also to identify its inner variations, to link the observed changes to biophysical and accessibility factors, determine the most influential environmentals factor and to predict quantity and spatiality of fore coming deforestation. There are four final research questions, which are later on specified with 8 sub-questions:

- **How were the forests and other land cover types distributed in Unguja spatially and quantitatively in 2009?**
- **How has the forest cover changed between 1996 and 2009?**
- **How have different environmental factors influenced forest changes spatially during this time period?**
- **How will the forest cover change in the future?**

2. Theoretical framework

2.1 Multiple values and uses of tropical forests

Forest and especially tropical forests sustain many important ecosystem services. Tropical forests cover only 7% of Earth's surface, while maintaining about the half of animal and plant species, making them primary terrestrial biodiversity hot spots. Species are driven to extinction and biodiversity is threatened when deforestation causes destruction, fragmentation or increased edge effect of habitats (Wilson 1988; Skole & Tucker 1993; Rudel & Roper 1997; Laurance 1999; Ferraz et al. 2003; Jongman & Pungetti 2004: 2–33). Agricultural, urban or other vegetated land covers bind significantly less carbon than forests, thus accelerating climate change (Skole & Tucker 1993; Laurance 1999; Houghton et al. 2000). Reduction of forest cover also increases the surface flow of watersheds, which exposes the land to flooding, water erosion and decreased infiltration and groundwater flow rates (Sahin & Hall 1994; Laurance 1999). Although deforestation might increase the water yield of lakes and rivers in a short run, it influences the hydrological cycle negatively causing less rainfall in a longer time period, due to decreased evapotranspiration (Skole & Tucker 1993; Laurance 1999; Maass et al. 2005; D'Almeida et al. 2007). Besides water erosion, cleared soils are also sensitive to wind erosion. Erosion along with the decrease of microbial biomass, loss of labile organic matter and disturbance of macroaggregates caused by deforestation, leads to decline in overall soil quality (Islam & Weil 2000).

Deforestation has also direct human-related effects. Forests are essential for livelihoods, especially in the developing world. They provide fuel wood, charcoal, medicinal plants, construction materials, wild fruits, beverages, spices, hunting grounds, materials for handicrafts, grazing areas for livestock and land for shifting cultivation for individuals and communities (Maass et al. 2005; Fagerholm & Käyhkö 2009; Swetnam et al. 2011; Fagerholm et al. 2012). In national scale, timber, minerals and ecotourism are essential sources of income and possibilities of economic development for many tropical developing countries (Laurance 1999; Maass et al. 2005). Also large amount of medicinal plants can be found from tropical forest. However only a small portion of these are scientifically tested and even smaller fraction are refined as drugs (Balick & Mendelsohn, 1992; Laurance 1999). Besides the direct material values many indigenous cultures are closely related to the used lands and forests and the knowledge collected by them is threatened as these are threatened (Alcorn 1993; Laurance 1999; Contreras-Hermosilla 2000). Even though culture would not be strongly built around the forest, these still have an important non-material value as sites for recreation, traditional believes, aesthetic and intrinsic values (Maass et al. 2005; Sitari 2005; Fagerholm et al. 2012).

2.2. Global and regional trends of forest change

Forests cover proximately one third of Earth's, but 0,13% (5,2 million hectares) of these forests are disappearing annually. The pace of net deforestation has lowered significantly since 1990s, when it was still 8 million hectares per year, but nonetheless the current development is unsustainable. All of the tropical regions are under this threat, but especially severe the situation is in South America and Africa. Although South America is often considered as the primary deforestation hotspot, Africa is not far behind. If reforestation is left acknowledged approximately 4 million hectares of forests are lost in South America annually, while the figure is only 0,6 million hectares less in Africa (FAO 2010; 11–22). Besides these figures the situations are completely different. In South America half of all land is forested and 80% of these are old growth primary forests, while in Africa only less than a quarter of the total area is under forest cover and only 10% of these are primary (FAO 2010: 15–34). In Eastern Africa the share of primary forests is even as lower (2,4%) (FAO 2010: 52).

The forests of Zanzibar are part of East African Coastal Forests (EACF), extending from Somalia to Mozambique (Figure 1). EACFs are combinations of multiple habitats and not only formed from closed forests, but also from drier woodlands, thickets and scrublands (Burgess & Clarke 2000: 84–94). Spatially comprehensive periodical forest surveys are missing from the area, but some estimations of the total forest extent are made (Dallu 2004). Godoy et al. (2011) estimated the EACFs to cover 3583 km^2 of Tanzania in 2000, while Tabor et al. (2010) measured 4560 km^2 cover. The former measured also the cover for 2007 when it was 2737 km^2. Though, differences in class semantics, forest detection methods and delineations of the ecoregion make comparison of various estimations difficult (Olander 2008). It is estimated that today's EACF cover only 5–40% from their original extent and the continuing decrease of forest stock is a serious problem in the area. Besides deforestation, also fragmentation and degradation are devastating forest systems by dividing continuous forests to patches generally smaller than 5 km^2 (Burgess 2000b; Dallu 2004).

It has been estimated that approximately three quarters of the EACFs are under high or very-high deforestation threat, while the remaining one quarter is under institutional protection (Tabor et al. 2010; WWF Tanzania Country Office 2012: 70). The estimated annual deforestation rate for whole Tanzania in 2010 was 1,13%, which was the second highest in Africa (FAO 2010: 21). Surprisingly the annual deforestation rates for Coastal Forest in Tanzania were lower than for the whole country, although again the methodological differences make comparisons biased. In the study of Tabor et al. (2010) 371 km^2 (8,1%) of these forests disappeared between 1990 and 2000, making the annual forest cover decrease to be 0,8%. At the same time period Godoy et al. (2011) measured annual deforestation rate to be around 1%. In general the annual deforestation rate seems to be decreasing and between 2000 and 2007 it was anymore 0,4% (Godoy et al. 2011). Deforestation faced especially non-protected areas, which had 5,5 to 9 times higher deforestation rates than the protected areas (Tabor et al. 2010; Godoy et al. 2011). Also the areas near the coastline were more prone to deforestation, although majority of low-lying forest have already gone through extensive deforestation process in the history (Prins & Clarke 2007; Tabor et al. 2010; Swetnam 2011).

Figure 1. The extent of East African Coastal Forests in 2006 modified from TFCG (2006).

Besides the pace of deforestation also the driving forces behind the process vary significantly between and within continents, countries and regions (Geist & Lambin 2001: 1–2). In Africa deforestation is more often linked to subsistence agriculture, shifting cultivation, fuel wood extraction, population growth, increased accessibility, foreign depth, urbanization and land tenure insecurity, while in South America it relates more to road and settlement expansion, commercial logging, grazing, rapid market growth, industrialization, taxation and capital (Geist & Lambin 2001: 23–49; Rudel et al. 2009). Some have argued that large scale loggings, export oriented agriculture and bioenergy are becoming more important drivers of deforestation in Africa, but at present the deforestation process is still strongly subsistence agriculture driven (Rudel et al. 2009; DeFries et al. 2010; Fisher 2011). These substantial inter- and intraregional differences emphasizes the value of research done at various geographical levels, ranging from local to global (Lambin 2003).

It is estimated that 60% of the Coastal Forest have been converted to farmland or urban areas (WWF Tanzania Country Office 2012: 70). The forests are still threatened by urban and agricultural expansion, but also timber cutting and charcoal production are serious threats (Burgess et al. 2000; Ahrends et al. 2010; Tabor et al. 2010; Swetnam et al. 2011). WWF Tanzania Country Office (2012: 70) and Swetnam et al. (2011) made stakeholder estimations about the proximate causes of deforestation in EACFs. In these estimations agricultural expansion was seen as the number one cause of decline, but also demand for fuel wood was analyzed as very high threat. Infrastructure expansion, unsustainable logging and forest fires were seen also as influential driving forces, while harvesting wood for carving, grazing, unsustainable hunting, conversion to salt pans, mining, climate change, collection of materials for selling, invasive species and pollution were analyzed as minor threats.

Majority of the Coastal Forest in Tanzania are Forest Reserves in their management status, while approximately one-fifths of the area is under more strict conservation as National Parks or National Reserves (Dallu 2004; Tabor et al. 2010). However many areas lack management plans or they are badly outdated. Where plans exist, limited resources and high pressure makes difficult to achieve their objectives and although general rules for forest use are set, they are rarely enforced locally. Nationally, environmental protection is promoted, but in reality it is rarely integrated in to economic policies, leading to situations where environmental values are overruled by the economic ones (Dallu 2004). Because of the severe lack in administrative resources, Participatory and Joint Forest Management schemes promoting greater role of local communities are widely promoted in Tanzania (Dallu 2004; WWF Tanzania Country Office 2012: 68 – 69).

2.3. Forests and dynamics of forest change

FAO (2000) has defined *forest* as a land area that is larger than 0,5 hectares, where trees reach minimum height of 5 meters in maturity and canopies cover at least 10% of the area. Forests are further subdivided to natural and planted, former referring to indigenous forests and latter to planted ones. Even though an area would fill these land cover requirements it should not be considered as forest if the main active land use is agricultural or urban, thus agroforests are not considered as forests in FAO (2000) classifications. Based on the same standards *deforestation* is a situation where forest converts to other land uses or the canopy coverage decreases below 10%, due to anthropogenic or natural reasons. The conversion needs to be long-term or even permanent and therefore it should last at least 10 years (Kaimowitz & Angelsen 1998; FAO 2000; Rudel 2005: 12). Long-term fallows and secondary forests are often considered as forests, but in certain cases these may be left outside the concept (Kaimowitz & Angelsen 1998).

Forest changes are truly geographical phenomena influenced by both the social and natural realms of life. The drive to change often comes from comprehensive happenings in the society such as population growth, urbanization or increased market demand (Figure 2) (Kaimowitz & Angelsen 1998: 90–98; Angelsen 1999; Geist & Lambin 2001: 6–8). Environmental realm on the other creates the boundary conditions for this change. Forests simply cannot grow where the soil or climate is unsuitable and are rarely deforested from areas inaccessible to humans (Rudel & Roper 1997; Kaimowitz & Angelsen 1998: 90; Geist & Lambin 2001: 13–15; Verburg et al. 2004). The underlying social causes determine the quantity of change, while environmental factors designate its spatial pattern. Though, sometimes environmental factors can create feedback effects which work as the underlying causes (Geist et al. 2006). For example deforestation leading to soil degradation and unpredictable water flowes causes decrease in agricultural productivity. Decreasing productivity leads people to seek income from forest products, which eventually causes more deforestation. Similarly social underlying causes, like political decisions of forest conservation, have an influence on the spatial pattern of forest change.

Underlying causes and environmental boundary conditions come to reality through actions of individuals and communities towards forests. These actions, such as fuel wood collection and shifting cultivation are called as proximate causes and they always have a direct impact on forests (Kaimowitz & Angelsen 1998: 90–98; Angelsen 1999; Geist & Lambin 2001: 6–8). All of these direct changing processes have their own spatio-temporal nature, which is influenced by the underlying causes and environmental factors. For example fuel wood collection degrades forests steadily around the rural villages from where usable wood is available, while shifting cultivation spreads more widely to forested areas nearby, but causing only a short-term change (Geist & Lambin 2001: 69–71; Ahrends et al. 2010; Hett et al. 2012). Eventually it is the quantity and spatial pattern of the change process, which becomes empirically verifiable. Happened changes provide feedback to the underlying causes and environmental factors, which may alter the process when it starts again from the beginning (Geist & Lambin 2001: 14–15).

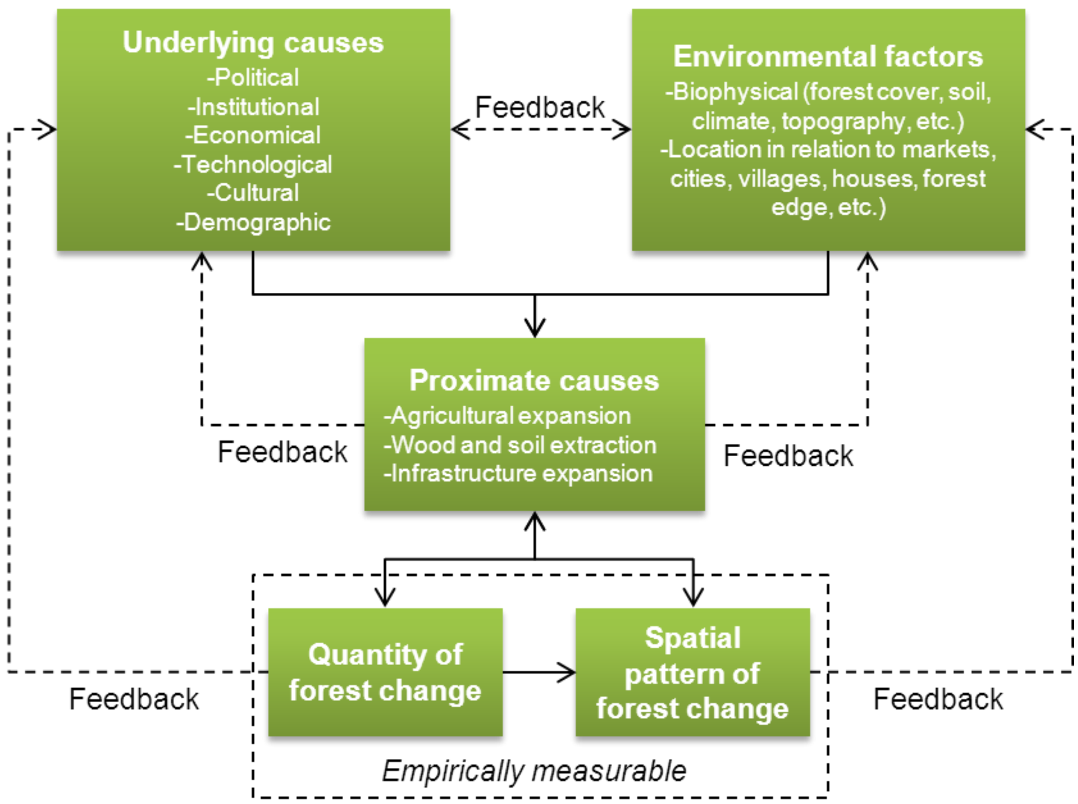


Figure 2. Dynamics of forest change. Underlying causes dictate the need for change, while environmental factors determine its boundary conditions. The changes are actualized through proximate causes, which all have their own spatio-temporal characteristics influenced by the underlying causes and environmental factors. It is eventually the quantity and spatial pattern of the process which becomes measurable.

As an example of the forest dynamics modified from the case study of Nagendra et al. (2003): Increase in coffee prices in the Western markets (underlying cause) promotes establishing new coffee plantations in Guatemala, but coffee only grows in high altitude and in particular soil conditions (environmental factors). All of the forests with the right altitude and soil could be deforested, but the coffee farmers only prefer those sites that are along main roads (environmental factor) and plantations concentrate diffusively there (proximate cause). Unfortunately the deforestation concentrates within the habitats of endangered cougars and after political pressure from the environment movement the government declares all cougar habitats protected (feedback). After the decision the new coffee plantations are abandoned and reforested (temporal nature of change). Eventually these actions lead to certain quantity and spatio-temporal patterns of forest change, which can be measured with remote sensing and GIS techniques. At the end of the day forest dynamics are created by countless similar and sometimes entangled processes, which alter the forests constantly, making it extremely difficult to fundamentally understand the process and the forces behind it (Contreras-Hermosilla 2000; Geist & Lambin 2001: 95–96)

2.4. Underlying causes

Mere detection of forest change is insufficient, if it is not connected to the forces causing it (Meyer & Turner 1992). Underlying causes are large social processes that are behind or underpin majority of the proximate causes (Table 1). These are political, social, economic, technological, demographic and cultural factors operating at various spatial scales, influencing forest change either quite directly at local level or indirectly through various chain-linked systems of national, regional or global scales. Eventually they are connected to proximate causes and actors at local level executing the actual change process (Contreras-Hermosilla 2000; Geist & Lambin 2001: 8–13; Lambin et al. 2001).

Table 1. Underlying causes of deforestation (Kaimowitz & Angelsen 1998: 95–98, Contreras-Hermosilla 2000; Geist & Lambin 2001: 8)

	Underlying causes
	Economic
Market growth	Rapid market growth, market failures, rise of cash economy, increased market accessibility, growth of certain industries, growth of consumer good demand
Economic structures	Poverty, unemployment, low living standards, economic downturn, external debt
Urbanization & industrialization	Growth of urban markets, increasing basic, heavy and forest based industries
Economic parameters	Advantages of cheap production and materials, price changes of certain commodities, price changes of cash crops
	Policy and institutional
Formal policies	Taxes, tariffs, subsidies, licenses, bans, finance, legislation, land use, zoning, transportation, forest policy, structural adjustment
Informal policies	Corruption, lawlessness, mismanagement
Property rights	Insecure ownership, land tenure, land competition, access to land
	Technological
Agro-technological change	Land use intensification and extensification, production changes
Technological applications in forestry	Damage due poor logging, wastage in production
Other production factors in agriculture	Low level of technological inputs, land scarcity, limited labor, lack of capital, lack of irrigation
	Cultural
Attitudes, beliefs and values	Lack of forest protection, low education, frontier mentality, nation-building, concern about welfare of future generations etc.
Individual and household behaviour	Increase in demand, consumption, commercialization, traditional use of resources
	Demographic
Demographic	Population pressures, population growth, migration, population density, population distribution, population structure, etc.

While connections between proximate causes and forests changes are quite straightforward, it is more difficult to affirm the causality between underlying forces and change (Kaimowitz & Angelsen 1998: 95–98). Underlying causes tend to operate in chain-linked manner where the same cause can have different influences at different locations at different times, depending on how it is linked to the rest of the cause-chain. This makes it difficult to lay universal explanations, but certain regional or national generalizations can be made (Geist & Lambin 2001: 1–2; Lambin et al. 2001; Turner et al. 2007). However the underlying causes are often outcomes of other underlying causes, making it hard to pinpoint one and only driving factor (Contreras-Hermosilla 2000; Geist & Lambin 2001: 95–97). For example poverty often drives people to unsustainable cultivation methods, making poverty the underlying cause of deforestation, but poverty can be also seen as outcome of population growth or unequal power over resources caused by colonization. So is it eventually poverty, population growth, unequal power relations or colonization, which is the main underlying cause behind deforestation? It would be easy to bypass these kinds of chain-linkages as causalities in history, irrelevant to actions taken against present deforestation, but to change the current state, underlying causes need to be responded, and the answers given to poverty are very different from the ones given to population growth or unequal power relations (Contreras-Hermosilla 2000).

Because underlying causes are extremely diverse and vary between regions, only those that have significant value in Africa, or in Tanzania, are introduced here. From the 19 deforestation case studies from Africa analyzed by Geist & Lambin (2001: 23–49) demographic factors were influential in 95%, economic factors in 84%, technological factors in 74%, institutional and policy factors in 47% and cultural and sociopolitical factors in 37%. Market growth, market failures, changes in agricultural prices, commercialization, growing population densities, in-migration, agrotechological changes and wood sector related matters were the more specific factors behind deforestation (Angelsen & Kaimowitz 1999; Geist & Lambin 2001: 23–49).

Geist & Lambin (2001: 46) found links between deforestation and such demographic factors as, population growth (in 79% of cases), in-migration (47%), population density (32%) and urbanization (26%). Already since Malthus *population growth* has been viewed as a central driving force to land cover change (Geist et al. 2006: 53). Growing population pressure influences deforestation in multiple ways. Directly it increases rural population seeking for agricultural land, fuel wood, timber and other forest products and the amount of customers for these products in-situ and in urban areas. More indirectly it affects the labor markets by pushing down wages, which causes people to seek incomes from agriculture and forest activities that were unprofitable earlier or through the institutional and technological changes (Kaimowitz & Angelsen 1998: 95–96; Mertens et al. 2000; Scrieciu 2006). The sheer number of population is no longer the only issue mattering, also population composition, fertility, distribution, urbanization, migration, household size and household live cycle have gained interest in research (Rindfuss et al. 2005: 361–363; Geist et al. 2006: 53–54; Lambin et al. 2006: 6). Especially *in-migration* has become important demographic factor in the globalized world where mortality and birth rates are stabilizing. From all of the demographic factors it is precisely in-migration to sparsely populated primary forest areas that has been proved to cause deforestation with absolute certainty (Rindfuss et al. 2005: 357; Geist et al. 2006: 54).

Writght and Muller-Landau (2006) have argued that urbanization could severely decrease the deforestation pressure caused by population growth and help to avoid mass extinction of tropical forest species. Urbanization eases the pressure of population growth as rural migrants leave their old fields abandoned and to reforest. Urban habitants are also more open-minded towards environmental friendly technology developments and their monetary incomes are higher. A higher income promotes importation of food and forest products and decreases the domestic demand (Contreras-Hermosilla 2000: 21). It also increases forest product prices at home, which could lead to tree planting in rural areas (Foster & Rosenzweig 2003; Rudel et al. 2005). Urbanization is however a double-edged sword and on the other hand it is related with increasing deforestation (Geist & Lambin 2001: 9; Rudel et al. 2009). When urbanization leads to increased personal incomes it also leads to higher consumption of agricultural products, which pushes the agricultural lands to spread. Although this may not happen at home it increases deforestation generally (Rudel et al. 2009). Also in Africa charcoal is often the most used cooking energy in the cities and its production chain requires more forest materials than using unprocessed woodfuel (DCCFF 2008; Ahrends 2010).

Arguments saying that majority of deforestation is caused by population growth combined with poverty through shifting cultivation practices has been proved as crude simplifications (Lambin et al. 2001). It is coming more and more obvious that demographic matters are always connected to other causes of culture, economy, technology, formal policies, accessibility and biophysical settings, and the causality between demographic factors and land cover changes are outcomes of these connections rather than sheer population changes (Kaimowitz & Angelsen 1998: 95–96; Rindfuss et al. 2004; Geist et al. 2006: 54). Evidence of population factors influences at micro-scales has been scarce, but at macro-scales the effect is clear (Rindfuss et al. 2004). Differences between regions make generalizations fruitless and underline the importance of study area knowledge. Also the causalities of demographic factors and other underlying causes suggest that concentrating solely on demographic problems is rather limited approach to deforestation (Geist et al. 2006: 53–54)

In the cases analyzed by Geist & Lambin (2001: 34) such economic and political factors were connected with deforestation as growth in agriculture- and wood-related industries (in 68% of cases), growth in demand of agricultural and forestry products (68%), increased market accessibility (26%), economic development (26%), land policies (26%) and foreign debt (21%). Surprisingly poverty was not strongly linked to the deforestation cases of Africa (Geist & Lambin 2001: 36). *Increasing market demand* of agricultural crops and forest products leads to agricultural intensification and spread of agriculture and pasture land to forest areas (Kaimowitz & Angelsen 1998: 91–96; McConell & Keys 2005). Increasing market opportunities “pull” farmers to produce beyond subsistence, creating *growth in agriculture- and wood-related industries*, while same opportunities “push” to consume more than subsistence requires. Governments can increase these “push” and “pull” factors with subsidies, agricultural policies, taxation, and improving access to markets with infrastructure development (Kaimowitz & Angelsen 1998: 91–96; Mertens et al. 2000). Globally demand for meat, vegetables and fruits are increasing and causing subsistence croplands to decline and market oriented production to increase (Geist et al. 2006: 48; Rudel et al. 2009). These global and local trends associated with agriculture promoting policies can cause *increase in agricultural prices*. Higher prices make agriculture more profitable and capital available for farmers, which can generate deforestation, while decreasing prices may leave previously used lands uncultivated and free for reforestation (Kaimowitz & Angelsen 1998: 90–97; Mertens et al. 2000; Scrieciu 2006).

Foreign debt is considered to create deforestation when governments seek to generate income to pay debts through extensive logging projects (Conteras-Hermosilla 2000: 13–14). Although 21% of the cases from Africa studied by Geist & Lambin (2001: 36) link deforestation with public debt, the relationship is unclear (Kaimowitz & Angelsen 1998: 96–97; Conteras-Hermosilla 2000: 13–14). Methodological difficulties has made the subject difficult to prove scientifically, however top decision makers from forest rich tropical countries have been certain that the linkages exist (Conteras-Hermosilla 2000: 13–14).

Open access or communal *land ownership policies* are generally considered to cause deforestation, especially if forest clearance gives rights to land ownership. This encourages people to clear more land than needed for subsistence purposes, if they can sell the products or the land for profit later on or squeeze unwanted neighbors out from the area (Kaimowitz & Angelsen 1998: 94). Also clearing forests prevents land to be claimed under official protection, which would limit land owners possibilities for income generation (Conteras-Hermosilla 2000: 16). Hardin (1968) has named the phenomenon of maximizing self-interest at the expense of long-term communal good in communal lands as “the tragedy of commons”.

Geist and Lambin (2001: 36) did not link deforestation with *poverty* in Africa in their meta-analyses; however other researchers have done so. In poverty plagued surroundings there are limited off-farm employment possibilities, capital, technology and access to markets, eventually leaving no other possibilities to improve livelihoods than to turn more forests to cultivation, to shorten fallow periods, to overgraze or to harvest forest products for income (McPeak & Barret 2001; Khan & Khan 2009). Another issue is the ecological marginalization of the poor and politically weak. They are often pushed to marginalized lands already in the past and these lands do not allow livelihood generation without overexploitation, which has caused it to be status quo instead of something avoided until the end (Khan & Khan 2009). Poverty is also said to cause conservation opposing attitudes, when livelihood dependencies on natural resources are high, but on the other hand the dependency on natural resources is also said to promote conservation and other environment protecting actions (Scherr 2000; Swinton et al. 2003; Khan & Khan 2009). Eventually poverty-forest or poverty-environment interactions are highly place-dependent and local economic, institutional, political, social and cultural context determines how they function (Khan & Khan 2009).

Increasing market demand of agricultural products and their prices often promotes deforestation. These happenings are often outcomes of *increased income* and *economic growth,* but their effects vary significantly between regions. In regions with vast forest cover like Amazon, Southeast Asia and central Africa economic uplift leads to deforestation when nations try to capitalize their natural resources, while in forest scarce regions like in East Africa increase in forest prices often leads to reforestation (Kaimowitz & Angelsen 1998: 92–96; Contreras-Hermosilla 2000: 19; Unruh et al. 2005). Many researchers believe in Environmental Kuznets Curve hypothesis, arguing that that after reaching certain threshold, improved economy decreases forest pressure and increases the forest cover, by creating better functioning government institutions, increasing governments capabilities for conservation, creating technological improvements, off-farm employment, urbanization, shifts towards petroleum based fuels, recreational forest uses, enhancing conservation atmosphere and decreasing dependency on agricultural and forest products (Kaimowitz & Angelsen 1998: 92–96; Contreras-Hermosilla 2000: 19; Rudel et al. 2005; Scrieciu 2006; Barbier et al. 2010). Though the evidence supporting the hypothesis has not been consistent, the level of economic and forest cover thresholds has been unspecified and the local social, political, institutional and cultural conditions have not been acknowledged (Conteras-Hermosilla 2000: 19; Scrieciu 2006; Barbier et al. 2010). When Environmental Kuznets Curve hypothesis is more directly associated with forests it is referred as Forest Transition Theory or more accurately as its "economic development path", but there is also another route towards Forest Transition, called as "forest scarcity path" (Rudel et al. 2005; Barbier et al. 2010). In this route the declining forest cover linked with continuously growing population leads to rising forest product prices, which encourages tree planting instead of other land uses. Governments can accelerate this path through forest planting schemes and by controlling forest product prices (Foster & Rosenzweig 2003; Rudel et al. 2005). However the same uncertainties related to consistent evidence, level of thresholds and local conditions makes it difficult to believe in universal generalization of "forest scarcity path". The first forest transition path is often related with European countries, while the second links to such Asian countries as Bangladesh and India (Foster & Rosenzweig 2003; Rudel et al. 2005). Eventually these approaches relating economic changes to forest cover cannot be fundamentally understood without linking them to theories of Von Thünen and to the incomes generated by other land use options (Barbier et al. 2010).

In his theory Von Thünens argues that land is occupied by the use that generates the highest income and changes happen when other land uses became more profitable than the original one. As in the case of deforestation and agricultural expansion, deforestation takes place when agriculture becomes more profitable than keeping the forest untouched. The moment this actually happens, is influenced by multiple factors, such as agricultural input and output prices, technology, transportation costs and biophysical conditions and majority of these factors are linked to distance from markets or cities (Chomitz & Gray 1996: Verburg et al. 2004; Angelsen 2007: 2–4). In homogenous and isolated world Von Thünens theory creates a conceptual model, where most profitable and labor intensive agricultural land uses concentrate to the vicinity of cities, while steadily turning to less managed forest land uses when distance increases and profitability declines (Figure 3) (Angelsen 2007: 2–4). In genuine world Von Thünen model has proved to explain the spread of agriculture and the use of forest products (Verdburg et al. 2004; Ahrends: 2010), though the model is a crude simplification criticized not to recognize land conversion costs, variations in soil, climate, vegetation, costs of different transportation means or changes in demands. Anyhow it is current still today and Von Thünen's approach has gained wider meaning as studying land uses in relation to location and income (Chomitz & Gray 1996: Angelsen 2007: 2–4).

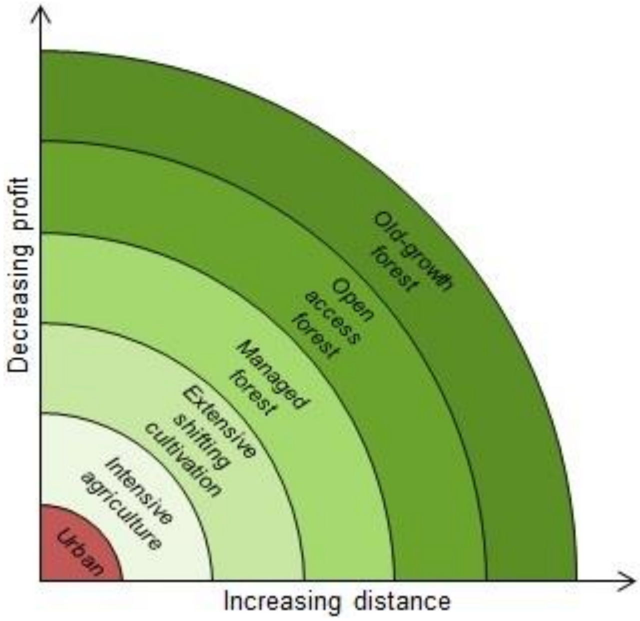


Figure 3. Land use based on Von Thünen model. Profits created by land uses decrease when distance from urban centers or markets increase, which influences the most prominent land use form in the area. Based on Angelsen (2007: 7).

2.5. Proximate causes

Proximate causes are those human actions that directly affect deforestation and they can be roughly divided to three categories: agricultural expansion, wood extraction and infrastructure expansion, which can be then subdivided to more precise causes (Geist & Lambin 2001: 6–7) (Table 2). Proximate drivers originate from land use decisions made at local level, which makes it easier to connect these factors to deforestation than the underlying causes. Underlying factors at national, regional or global levels can influence proximate drivers, but actual land use decisions are always made locally (Geist & Lambin 2001: 6; Geist et al. 2006; Turner et al. 2007). Some of them are more planned by the state or administration, like for example commercial logging actions or colonization of new areas, while some are more unplanned like intensified fuel wood extraction or shifting cultivation rotation (Angelsen & Kaimowitz 1999). Emphasizing only the proximate causes can lead to situations of "blaming the victim", as when subsistence farmers are accused for deforestation, although prevailing economic, political and technological situations may not allow them alternative options (Turner et al. 2007).

Table 2. Proximate causes of deforestation. Based on Geist & Lambin (2001: 7)

Proximate causes	
Agricultural expansion	Shifting cultivation Permanent cultivation Cattle ranching Colonization of agricultural land
Wood and mineral extraction	Commercial logging Fuelwood extraction Polewood extraction Charcoal production Lime making Coral extraction
Infrastructure extension	Transport infrastructure Market infrastructure Public services Settlement expansion Private bussiness infrastructure

From the African cases studied by Geist & Lambin (2001: 24–29) 84% linked deforestation with agricultural expansion, 68% with wood extraction and 47% with infrastructure extension. Though, rarely it is a single proximate cause behind the changes, but rather a combination of multiple causes. Generally deforestation in tropics and also in mainland Tanzania starts with logging of larger commercial trees, followed by shifting cultivation, collection of fuel wood and charcoal making in already degraded forest, finally leaving the area so degraded that complete shift to permanent agriculture becomes easy where soil conditions allow it (Lambin 1997; WWF Tanzania Country Office 2012: 73).

Agricultural expansion is seen as by far the most important direct cause of deforestation globally and in Tanzania (Kaimowitz & Angelsen 1998: 89; Geist & Lambin 2001: 85; Swetnam et al. 2011; WWF Tanzania Country Office 2012: 70). It is an umbrella for multiple change processes at various scales, ranging from individual farmers slash and burn actions to large scale agricultural colonization of tropical forests. In Africa agricultural expansion is linked mainly to subsistence and small holder agriculture (Geist & Lambin 2001: 25; Lambin et al. 2003). Shifting cultivation was seen as a driving force in 42% of the African cases analyzed by Geist and Lambin (2001: 25). Transmigration related agriculture was related to 21% of cases and cattle ranching to 16%. As said earlier, forests are not only threatened by the subsistence food crops, but also the market oriented cash crops have had their influence (Mertens et al. 2000; Geist et al. 2006: 48). Coconut, sisal, cashew, rubber and clove have placed areas previously occupied by forests and nowadays also such biofuel crops as jatropha, palm oil and sugar cane are threating the Coastal Forests (WWF Tanzanian Country Office 2012: 70–71).

Agricultural expansions have a general tendency to occur in low-lying and plain areas that have suitable soils and high water availability, located close to villages, cities and markets of agricultural products. Though, population pressure, soil degradation and subsistence needs can push farmers to cultivate areas unable to fill these requirements, which in tropics often leads to extensive shifting cultivation (Kaimowitz & Angelsen 1998: 90; Geist & Lambin 2001: 78). This is also the case in Coastal Forest – areas where soils are not widely able to support permanent agriculture. The high population pressure leads to overuse of agricultural land causing soil degradation and decreases in yields (Chidumayo 1987; Dallu 2004). This pushes agriculture to undisturbed forest areas and also intensifies the shifting cultivation rotation speed, leaving less time for the forest to regenerate and making them not only degrade horizontally, but also vertically (Chidumayo 1987; Hett et al. 2012). When possibilities to expand cultivation areas are limited, higher land use pressure appears mainly as shortening of fallow periods.

Agricultural expansion can form many different spatial patterns of deforestation. Small holder and subsistence based shifting cultivation creates a diffusive structure where patches spread across the landscape rather randomly, but within a close distance to home villages, while permanent agriculture creates more of a patchy structure, where forests are cleared almost completely leaving only small patches standing in areas unsuitable for agriculture (Figure 4) (Lambin 1997; Geist & Lambin 2001: 69–71; Hett et al. 2012). However the created structure depends heavily on the underlying factors and for example colonialist agricultural expansion often follows roads in corridor or fishbone type manner and if rural living is not nucleated to villages, individual households may cause island pattern of deforestation (Skole & Tucker 1993; Geist & Lambin 2001: 69–71).

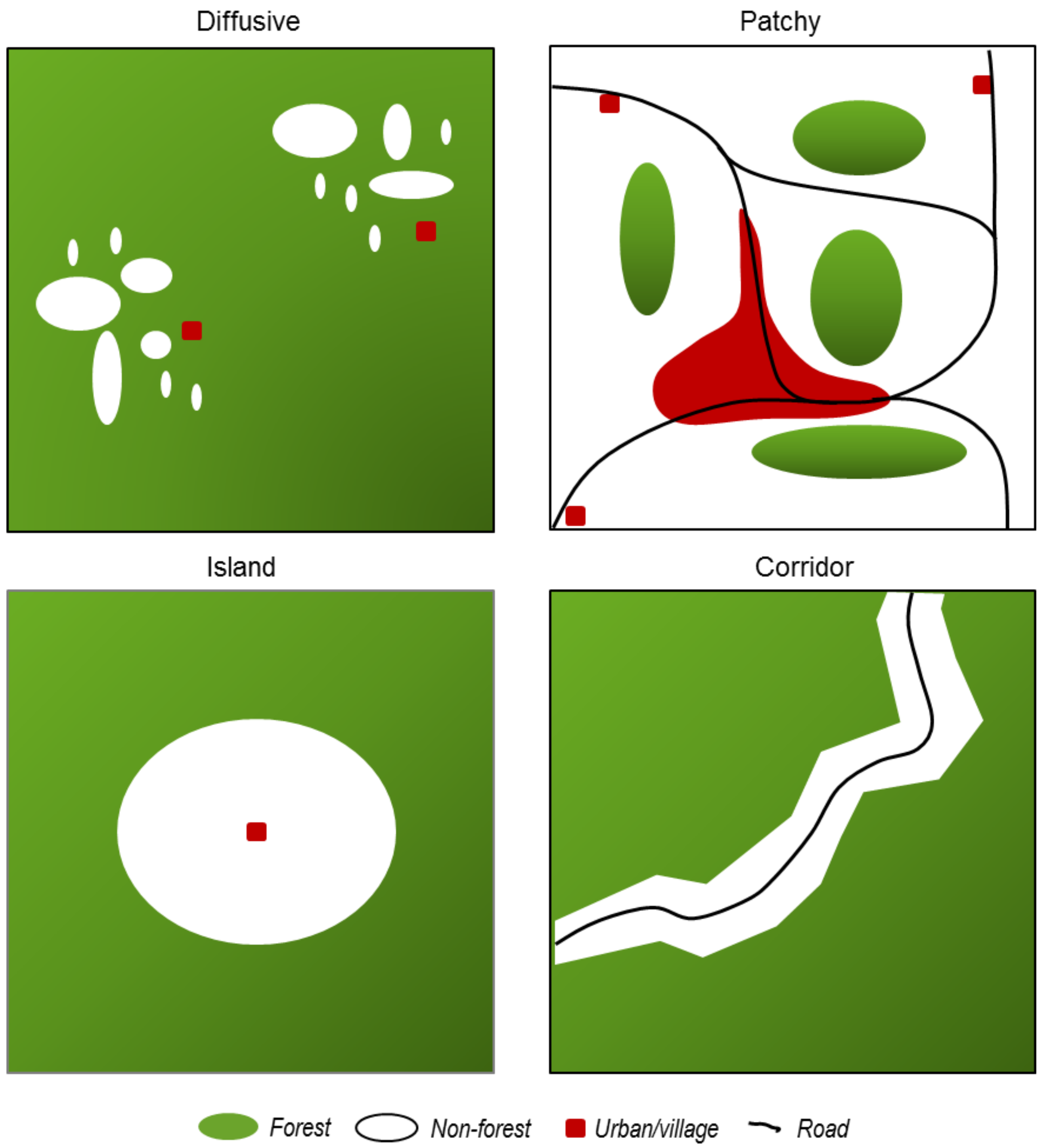


Figure 4. Different morphologies or spatial patterns of deforestation based on Lambin (1997) and Geist & Lambin (2001: 69–71).

While deforestation related to wood extraction is mainly due to commercial logging in Asia and Latin America, in Africa it relates more to fuel wood collection, charcoal production and polewood extraction (Contreras-Hermosilla 2000; Geist & Lambin 2001: 29). In study of Geist & Lambin (2001: 29) fire wood collection and charcoal making were behind over half of the deforestation cases from Africa, while polewood extraction influenced 42% and commercial logging only 26% of the cases. In Zanzibar also lime making and extraction of coral for construction purposes are deteorating the forest cover (Orjala 2008; Fagerholm et al. 2012). Charcoal is used mainly in larger cities and its influence is directly relational to accessibility to these cities, while fuel woods and polewoods are used more in rural areas, making their influence to spread more steadily

(Ahrends et al. 2010; WWF Tanzanian Country Office 2012: 71–75). As an example, charcoal provided to Dar es Salaam is estimated to influence areas as far as 300 km from the city and the production radius has spread over 50 kilometers since 1970s (Ahrends et al. 2010).

Compared to large scale loggings wood fuel collection, polewood cutting and other domestic wood extraction activities may have less significant deforestation effects at regional scale, but their influence is strong at local or village level (Geist & Lambin 2001: 72; Ahrends et al. 2010). Wood extraction for domestic purposes has been associated with island pattern of forest clearance, where deforestation happens mostly in areas circularly surrounding villages. Wood is simply extracted from areas which are closest and deforestation is not influenced by roads, soils, markets or topography, if they do not directly affect the availability of wood (Mertens & Lambin 1997; Geist & Lambin 2001: 71–72; Ahrends et al. 2010). Distribution of forest clearance caused by commercial logging, charcoal making and polewood extraction is dominated by vegetation in forest scarce regions, because these activities require relative large tree stems. Deforestation caused by selective logging of larger individual trees or unsystematic collection of fuel wood are difficult to measure with medium resolution satellite images, where minor changes do not cause shifts from forest to other land cover class, but rather cause vertical and qualitative degradation of forest stands (Healey et al. 2007: 68).

Infrastructure causes direct deforestation when forests are cleared for urbanization, roads, railroads, mining or other human facilities (Geist & Lambin 2001: 7). At global scale direct deforestation impacts of infrastructure expansion are minor compared to wood extraction and agriculture. In global land cover estimations dense built-up areas cover less than 0,2% of terrestrial land, while the share is around 20% for agriculture. Although built-up areas are underestimated because of the large pixel size used and the elongated shape of built-up areas, it is clear that infrastructure expansion has rather limited direct impacts on deforestation (Hansen et al. 2000; Loveland et al. 2000; Geist & Lambin 2001: 8). However indirect impacts are notable. Urban areas expand to close-by fields pushing agriculture towards forest, growing demands of urban population degrade forests nearby and transportation infrastructure opens access to new forest areas and links these to markets (Chomitz & Gray 1996; Mertens & Lambin 1997; Lambin et al. 2003; Verburg et al. 2004; Ahrends et al. 2010).

All African case studies analyzed by Geist and Lambin (2001: 27) linked deforestation with road extensions and only in couple of cases development of railroads and accessibility to settlements and markets had an important influence. Generally infrastructure extensions causes patchy or corridor type distribution of deforestation. Patchy distribution is linked to overall spread of urban and agricultural land covers, while corridor type deforestation refers to process where forests are cleared mainly from road sides leaving rest of the area unaffected (Geist & Lambin 2001: 68–71).

2.6. Environmental factors

Environmental factors rarely trigger deforestation processes, but merely shape and provide feedback to them, making them to drive the location, not the quantity of change. As said earlier, permanent agriculture does not spread to soils unsuitable for farming or infrastructure is not built to rugged slopes (Mertens & Lambin 1997; Rudel & Roper 1997; Kaimowitz & Angelsen 1998: 90; Geist & Lambin 2001: 13–15; Geist et al. 2006: 44; Verburg et al. 2004). Sometimes environmental factors can create feedback effects which work as proximate causes to the quantity of deforestation (Geist et al. 2006). Environmental factors can be subdivided to biophysical and location properties. Biophysical properties include soil, climate, topography, hydrology and vegetation, while locational properties are more related to human activity and are turned to such variables as accessibility from coastline, vicinity to roads and accessibility to markets or urban centers (Geist & Lambin 2001: 13–15; Geist et al. 2006: 44; Verburg et al. 2004). Locational variables connect pixels to society, while biophysical properties affect land use decisions on the spot (Irwin & Geoghagen 2001; Verburg & Veldkamp 2003). Different proximate causes can be linked to different environmental factors and vice versa (Figure 5). However quantitatively arguing anything related to proximate or underlying causes based on research connecting deforestation to environmental factors is methodologically problematic, since all the proximate causes are linked to many different environmental factors and all the environmental facors are linked to many diffent proximate causes (Veldkamp & Lambin 2001).

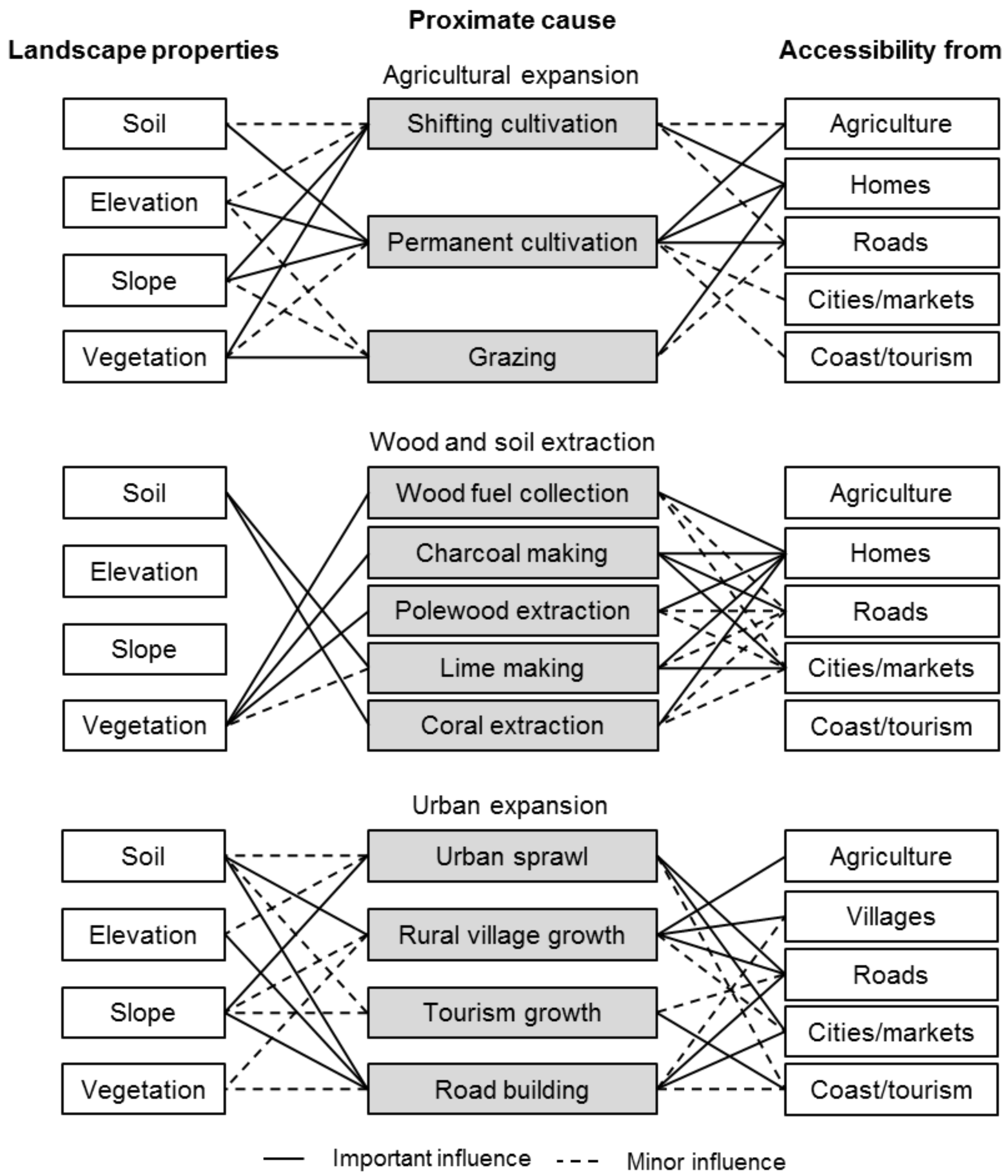


Figure 5. Causalities of environmental factors and proximate causes of deforestation adapted to conditions of Coastal Forests of Tanzania and Zanzibar based on related research (Chomitz & Gray 1996; ; Mertens & Lambin 1997; Angelsen & Kaimowitz 1999; Geoghegan et al. 2001; Verburg et al. 2004; Angelsen 2007: 2–4; Prins & Clarke 2007; Fagerholm & Käyhkö 2009; Tabor et al. 2010; Swetnam 2011; Fagerholm et al. 2012;).

Surprisingly many land cover change studies do not use *land cover* as a variable explaining the spatial pattern of deforestations. Different land cover types have different path dependencies altering the possibilities of changes. Fallows have clearly higher probability to reforest than urban areas and field cannot change back to primary forest. These limitations should be acknowledged in any land cover change modeling (Verburg & Velkamp 2003; Verburg et al. 2004). Land cover and its patterns also influences deforestation indirectly through the accessibility. It is easier to pass through open field than wetland to collect fuel wood and it is easier to collect these from edges of fragmented forests (Kaimowitz & Angelsen 1998: 93; Verburg et al. 2004). However land cover classifications are often crude simplifications with only few classes and no inner variations within the classes. These inner variations such as *vegetation* differences within forests may impact deforestation in multiple ways as different deforestation processes require different kinds of forest vegetation (Verburg et al. 2004). For example shifting cultivation driven deforestation tends to concentrate to areas of early secondary regrowth already used in cultivation, branches from only certain tree species are collected as fuel wood, while actions like logging, charcoal making and polewood extraction require trees to be relatively large and mature forests are more easily protected than low-lying scrubs or thickets (Geoghagen et al. 2001; Orjala 2008). Importance of vegetation is essential and Boonyanuphap (2005) found it actually to be the most important environmental factor explaining the spatial distribution of deforestation.

Fertile *soils* are preferred in agriculture and therefore have a higher tendency to deforest (Chomitz & Gray 1996; Irwin & Geoghagen 2001; Kok & Veldkamp 2001; Nagendra 2003). 50% of lands with fertile soils were under agricultural production in Belize, while this was only 15% for areas with less fertile soils (Chomitz & Gray 1996). However, when soil fertility differences are minor or fertilizers widely available, soil has significantly lower influence on deforestation (Serneels & Lambin 2001). Soils affect deforestation primarily through agricultural expansion, but they also have an influence on construction of roads, buildings and other infrastructure elements and can also be sources of geological materials such as sand or coral stones (Kaimowitz & Angelsen 1998: 94; Verburg et al. 2004; Fagerholm et al. 2012).

Elevation is important factor molding land use especially in areas with high altitude differences (Kok & Veldkamp 2001). Human population is more concentrated to low-lying coastal areas, which make these more easily turned to agriculture (Serneels & Lambin 2001). Higher elevation is also often connected to weaker accessibility, rugged terrain, poor soils and lower soil moisture, all relieving deforestation pressure (Geoghegan et al. 2001; Kok & Veldkamp 2001; Verburg & Veldkamp 2003). Although in some cases low-lying coasts are so dry, that areas slightly elevated are more suitable for agriculture and as an example coffee require shigher elevations to be grown in tropics, which leads to deforestation in these areas (Southworth et al. 2002; Nagendra et al. 2003). It is not only elevation which influences, but also topography in general. If terrain is undulating heavily or has sharp ridges these might make it undesirable for cultivation or impossible for infrastructure construction, which leaves it under less pressure, allowing unprofitable forest uses to bloom (Nagendra et al. 2003).

However, after forests are cleared elsewhere these rugged inaccessible patches might be the only proper forests left, making them the only possible areas for deforestation. If this is the cases, analyses connecting environmental variables with deforestation would show that elevated, rugged and undulating areas are more prone to deforestation, but proper consideration of historical context should be made before doing these kinds of wrong generalizations (Geoghagen et al. 2001).

Location of certain land or socio-economic elements can be turned to variables through accessibility measures. Accessibility refers to how well certain area can be accessed. It is not only influenced by distance, but also land cover types, road, rail road, river and airport networks, means of transportation, gas prices and topography (Kaimowitz & Angelsen 1998: 93). Accessibility is considered as one of the most important factors explaining deforestation patterns and different accessibility measurements have been used as independent variables in many spatial deforestation models (Ludeke et al. 1990; Chomitz & Gray 1996; Mertens & Lambin 1997; Angelsen & Kaimowitz 1999; Geoghegan et al. 2001; Kok & Veldkamp 2001; Serneels & Lambin 2001; Southworth 2002; Verburg et al. 2002, 2004; Nagendra et al. 2003; Aguiar 2007). Improved accessibility affects both the origin and the destination (Verburg et al. 2004). From origins perspective it allows access to new land and forest resources or lowers the cost of using old ones, while from destinations perspective it links these areas and their products to markets and innovations (Chomitz & Gray 1996; Irwin & Geoghagen 2001; Nagendra et al. 2003, 2004). In a sense, accessibility measures connect localities to larger scale phenomena’s (Verburg & Veldkamp 2002). It is most frequently measured from such sources as roads, railroads, waterways, coast, fields, forest edges, villages, cities, houses and markets with multiple means of transportation (foot, car, logging truck, boat) and with various measurement techniques (Euclidean distance, travel time, monetary cost, population potential) (Ludeke et al. 1990; Chomitz & Gray 1996; Mertens & Lambin 1997; Mertens et al. 2000; Geoghegan et al. 2001; Kok & Veldkamp 2001; Serneels & Lambin 2001; Southworth 2002; Verburg et al. 2002, 2004; Nagendra et al. 2003; Aguiar 2007).

Road accessibility is one of the most often used accessibility measures and proved to have significant influence on deforestation (Chomitz & Gray 1996; Geoghegan et al. 2001; Boonyanuphap 2005). Use of road accessibility relies on the idea that roads connect areas to markets, bring in innovations and in-migration, cause deforestation and fragmentation in building process and that their sides are more prone to spread of urban or agricultural expansions (Chomitz & Gray 1996; Geoghagen et al. 2001; Hett et al. 2012). Although accessibility through road network clearly influences deforestation, it seems that it is explain deforestation at local level, while correlations ca not be found at regional or national scales The effect of roads is also relational to their type and means of transportation available (Verburg et al. 2004; Boonyanuphap 2005). Altogether the causality between road accessibility and deforestation could be overestimated and not as straightforward as often presumed, since location of transport infrastructure correlates with population, land cover and geomorphology. Roads are built in areas already settled and cleared from forests and their alignment is often dictated by topography, soil quality and other biophysical factors. Also some accessibility measures that connect population with road networks, like population potential, make it impossible to separate the effects of population and roads (Kaimowitz & Angelsen 1998: 94; Verburg et al. 2004). Eventually influence of roads is highly chain-linked to other environmental and underlying factors. For example influence is higher when the soils along the roads allow permanent cultivation or when roads are built for colonialist settlements (Chomitz & Gray 1996; Contreras-Hermosilla 2000).

Market accessibility is considered as one of the most significant forms of accessibility influencing landscapes. Markets connect people and natural resources creating more demand and production, increasing forest clearance (Mertens et al. 2000; Geoghagen 2001; Nagendra 2004; Verburg et al. 2004). As was theorized already by Von Thünen in 1826, accessibility or distance to markets lowers the transportation causes of agricultural and other products, eventually increasing the lands rent, which promotes agricultural or urban uses instead of relatively unprofitable forest uses (Serneels & Lambin 2001; Angelsen 2007: 2–4). Von Thünen links markets with *urban centers*, which are still essential places for global and regional markets, but besides these also transport networks, hubs, harbors and airports bring products to the wider world, while some products are used locally (Verburg & Veldkamp 2003; Nagendra 2004). Though, often lack of data and efficiency drives to combine market and urban accessibility measurements as one variable (Irwin & Geoghagen 2001). Urban centers influence through markets but also by spreading infrastructure to their fringes and by pushing agriculture areas further (Lambin et al. 2003; Antrop 2006a). The correlation with deforestation is diverse and non-linear, since areas closest to urban centers are often already cleared from forests and deforestation occurs only after certain threshold (Mertens & Lambin 1997).

Accessibility to homes, villages and *population* is able to present pressure to the nearby lands where majority of agriculture and collection of forest resources takes place (Ludeke et al. 1990; Mertens & Lambin 1997; Geist & Lambin 2001: 69–71; Geoghagen et al. 2001; Fagerholm & Käyhkö 2009; Fagerholm et al. 2012). According to Verburg et al. (2004) the time distance between farmers home and fields is more important than the distance to markets, because this journey is taken almost daily, while products are taken to markets only occasionally. Also some products may be processed at the villages before transportation to markets or used locally (Verburg et al. 2004). In village case studies from Zanzibar, subsistence uses of land were averagely located 500 to 1200 meters from inhabitants' homes, livestock keeping and collection of construction materials concentrating more close and cultivation and fuel wood gathering more further away (Fagerholm & Käyhkö 2009; Fagerholm et al. 2012).

Based on spatial autocorrelation and diffusion theories agriculture has a tendency to concentrate near to areas already cleared for cultivation, thus *vicinity* or *extent of agricultural* areas is often used to explain spatial progress of deforestation (Geoghagen et al. 2001; Serneels & Lambin 2001; Aguiar et al. 2007). Various individual or household level decisions may cause deforestation near to already established agriculture: farmers may decide to expand their fields, swiddening causes spatio-temporal mosaic of fields, scrubs and forests and newly migrated farmers favor areas near old fields where the productivity of land is already tested. In geographically large swiddening landscapes deforestation tends not to happen near active agriculture, but rather in forest regrowth areas, while in other surroundings distance from other agriculture was one of the most important environmental variables predicting future changes (Geoghagen et al. 2001; Boonyanuphap 2005). The influence of nearby agriculture is greater in areas of permanent cultivation on fertile soils, while the pattern and linkages are more random at shifting cultivation landscapes (Kaimowitz & Angelsen 1998: 93). Increased accessibility can also lead to agricultural intensification instead of deforestation if there is a severe underuse of the current fields (Verburg et al. 2004).

Generally the areas deep within the forest are safer in terms of deforestation than the areas on the edges where majority of pressure occurs. Highly fragmented forests having more edge are therefore more easily deforested (Ludeke et al. 1990; Kaimowitz & Angelsen 1998: 93; Rudel & Roper 1997; Fox et al. 2003; Nagendra et al. 2003, 2004). *Distance from forest* edge is often used variable and it has proved influential in predicting future changes (Ludeke et al. 1990; Mertens & Lambin 1997; Nagendra et al. 2003). Even though certain change processes, as selective logging or forest fires, may alter this generalization (Healey et al. 2006; Clark & Bobbe 2006). In case study from Cameroon, Mertens and Lambin (1997) showed that 80% of deforestation happened within one kilometer from forest edge and it was significantly better explanatory factor in shifting cultivation environments, than the vicinity of roads.

In island surroundings *distance to coast* has its own influence. Coastal and island forest are more easily accessible than large unifies patches in continental regions (Kaimowitz & Angelsen 1998: 93). In East Afrcian Coastal Forest these areas are largely already turned to more profitable agricultural uses, but remaining patches are under heavy pressure (Prins & Clarke 2007; Tabor et al. 2010; Swetnam 2011). Tourism creates its own forces of change, when coastal and especially beach locations are favored by hotel constructors and building materials are collected from nearby areas (Käyhkö et al. 2011).

3. Methodological framework

3.1. Land change science as an approach for deforestation research

The causes, spatial pattern and consequences of deforestation are extremely multidimensional and linked to various societal and environmental processes. Understanding the nature of deforestation requires a combination of scientific approaches from social, environmental and geographical sciences, such as landscape ecology, biogeography, political ecology and demography joined with methodologies of remote sensing, GIS and statistics. A true combination of these is referred as Land Change Science (LCS) (Rindfuss et al. 2004; Turner et al. 2007). LCS seeks answers to questions related to land change observation, monitoring, causes, consequences, modeling, vulnerability, resilience and sustainability, studied at scales ranging from local to global (Rindfuss et al. 2004; Turner et al. 2007). Complexity of human-environment and space-time relations, combination of approaches from multiple disciplines and diversity of land changes create their own challenges for LCS (Rindfuss et al. 2004). Each change is an outcome of multiple agents, multiple uses of land, multiple responses to social, economic and environmental contexts, multiple spatial and temporal scales and multiple connections between people and their land. This makes building universal theories extremely demanding, but complicated study subjects such as climate change and biodiversity loss call for more overarching theories and increasing amount of articles, journals and empirical evidence are leading the way towards this (Lambin et al. 2006: 7; Turner et al. 2007).

LCS is not ready to produce these overarching theories, but the matters they should relate to, are laid. Firstly the land change theory should be able to lay some rules how land units (ea. forest pixel, urban patch) change, how social units (household, communities, administration) cause changes and how environmental factors (ea. soil, climate, elevation, slope and location) distribute these changes spatially. These generalizations can be given based on the environmental and historical settings (soil, climate, elevation, slope, location and change history) of the land unit, and by socio-economic factors (age, wealth, education, household life cycle and livelihoods) of the social unit (Lambin 1997; Lambin et al. 2006: 8). Secondly the theory should acknowledge the importance and linkages of scale. Social units are always combinations; individual belongs to a household, household to a village, village to region. In some cases it might be the factor working at household level causing the changes, while sometimes global economic shift can be the main driving force. Similarly also the land units belong to larger groups, such as continues urban patch, shifting cultivation area or watershed, all having their own functionality (Lambin et al. 2006: 8). Therefore it should always be considered that which social unit answers land changes of that particular land unit at focus (Rindfuss et al. 2004; Turner et al. 2007). Thirdly, the theory should be able to link the past with the present and the future. Historical context of both social and land units determines their change possibilities in the future. For example communities that have relied on fishing for generations are less eager to cause deforestation through agricultural expansion than migrant households with agricultural backgrounds. For land units there are certain path dependencies existing: urbanized areas do not reforest without major disasters and primary forest are not as vulnerable to deforestation as secondary regrowth (Lambin et al. 2006: 8).

Land cover classifications and change detections work as the starting points for any deforestation or LCS research (Ridfuss et al. 2004). Though the subtleties of these approaches have developed enormously in recent decades, the focus of the LCS community has shifted from observing the change to explaining it. Local and regional level spatial patterns are explained by connecting change detection outcomes to environmental factors with distance analyses and spatial regression models in GIS (Ludeke et al. 1990; Mertens & Lambin 1997; Nagendra et al. 2003; Verburg et al. 2004). Proximate and underlying causes are linked to the processes with somewhat more aspatial regression methods in statistical softwares or with qualitative or narrative analyses and used to explain national and global differences (Geist & Lambin 2001: 18; Verburg et al. 2002: 118–119). These approaches provide knowledge about the quantity, spatial extent, key areas and causes of deforestation for the stakeholders, decision makers and global scientific community. However, also more forward looking approaches have been developed to help the decision makers to allocate actions in spatially more accurate manner (Verburg et al. 2004b). One of these developments has been to predict deforestation spatially by projecting the outcomes of change detections and spatial regression analyses in to the future (Mertens & Lambin 1997; Veldkamp & Lambin 2001; Alcamo et al. 2002; Verburg et al. 2004b).

3.2. Remotely sensed data and its classification to land covers

Analyzing spatial distribution of forests and other land cover types requires a land cover classification and detecting changes requires at least two of these snapshots of time. If appropriate or updated classifications do not exist, classifications can be created by reclassifying of old maps or with such remote sensing methods as manual digitization of aerial images or automated classification of satellite images (Lillesand et al. 2008: 189–222, 545–591). Remote sensing as such is based on measuring and analyzing the electromagnetic reflectance from the ground. More precisely optical remote sensing measures the electromagnetic reflectance of Sun from Earth's surface at the wavelengths of visible light, near- and mid-infrared (0,4 – 3,0 µm) (Lillesand & Kiefer 1994: 6–9). Different land cover features have different combinations of reflectance values at different wavelengths and this allows separating them (Figure 6) (Campbell 1996: 314–315; Lillesand et al. 2008; 545–547).

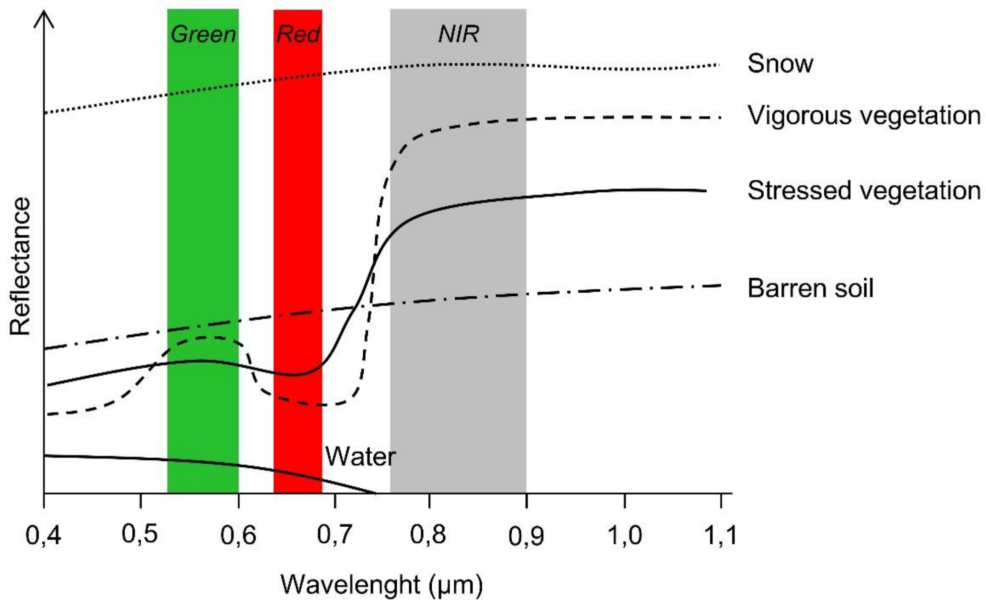


Figure 6. Reflectance of different land cover features at different wavelengths. The color bands in the middle represent different multispectral bands (red, green and near infra-red) of the Landsat TM image. Modified from Lillesand et al. 2008: 17.

The electromagnetic reflectance is captured by remote sensing sensors, which turns it to an image. However the journey from Sun through Earth to a sensor is long and contains many uncertainties. Reflectance can be scattered, absorbed or bended while traveling through the atmosphere, which is a severe problem, especially for the satellite sensors located far away (Campbell 1996: 30; Lillesand et al. 2008: 9–12). Also different surfaces reflect radiance differently. There are no perfectly straight surfaces in nature as there neither is perfectly scattering surfaces, therefore all of the radiance reflected by an object never reaches the sensor and the objects in remotely sensed images always contain radiance from other objects (Campbell 1996: 30; Lillesand et al.

2008: 12–29). Also topography, angle of the sensor and azimuth of the Sun influences the amount of radiance collected by the sensor (Mather 2005: 18; Lillesand et al. 2008: 12–29).

The sensor captures the electromagnetic reflectance and turns it to an image. Although the images seem to be continuous representations, they are actually created by combination of predefined two-dimensional square shaped individual picture elements, called pixels. Each pixel has certain brightness value, which corresponds to the average radiance value measured electronically from its geographical area. In multispectral images there are multiple bands representing different wavelengths of light. The average radiance value at each partition of the wavelength spectrum is recorded to different bands (Lillesand et al. 2008: 30–33). The absolute radiance values are transformed to more computation friendly Digital Number (DN) values, ranging from 0 to 255, 0 to 511, 0 to 1023 or higher. These ranges are defined by the binary computer coding scales called 8-, 9- or 10-bit, respectively. In these formats the image data can be automatically visualized by the computer (Campbell 1996: 314–315; Lillesand et al. 2008: 30–33, 545–547). So eventually the radiance values collected for example by the Landsat TM sensor are turned to brightness values ranging from 0 to 255 measured separately from seven different parts of the wavelength spectrum and visualized as pixel of 30 m^2.

The remote sensing sensors capture reflectance differently. They have unique *spatial resolutions*, referring to the geographical area represented by individual pixels (Table 3). This limits the size of the smallest unit able to be detected. In the modern high-resolution satellite sensors the spatial resolution is less than 1 m, while in Landsat TM it is 30 and even 1000 m in such sensors as MODIS and AVHRR (Mather 2005: 34–48; Lillesand et al. 2008: 30–36). The larger the pixel size is the more there are “mixels”, pixels which are representations of multiple objects instead of purely representing one (Foody 2004). However the larger pixel size reduces the data processing times and allows collection of globally continues land cover data. Besides the spatial resolution, sensors also have also differences in their spectral, radiometric and temporal resolutions. Different sensors measure different partitions of the electromagnetic spectrum and collect it to different amounts of spectral bands. The used spectrum is referred as spectral resolution. The temporal resolution refers to the timespan between images taken from same area and the radiometric resolution to the used binary system. Besides these also the coverage of sensors varies. In general the coverage is rather small in high-resolution satellite sensors, while low-resolution sensors cover large areas at once (Lillesand et al. 2008: 30–36).

Table 3. Differences in spatial, temporal and spectral resolutions between satellite remote sensing sensors. Source: Mather 2005: 34–48; Lillesand et al. 2008: 399–481.

Satellite	Sensor	Spatial resolution (m)	Coverage (km)	Temporal resolution (d)	Spectral resolution (µm)
Geoeye-1	HRV	0,41–1,65	15,2	>3	0,45–0,90
IKONOS-2	OSA/TDI	1–4	11	2–3,5	0,45–0,90
SPOT-3	HRV	5–20	60	26	0,49–0,89
Landsat 4	TM	30–120	185	16	0,45–2,35
NOAA-17	AVHRR	1000	3000	0,5	0,55–12,5

Although there are various sensors existing, Landsat imaginary have become the backbone of land change research. The satellite has provided continuous Earth Observation data since 1972 and its creators have committed to continue collecting "Landsat-like" imaginary also in the future (Lillesand. et al. 2008: 399–432). The long history of observations, medium sized resolution, relatively large coverage, good spectral resolution and relatively short temporal coverage have all made Landsat extremely popular, but extensive cloud coverage and lack of acquiring, storage and distributing facilities in parts of the tropics have undermined its continuity. However these data gaps can be occasionally filled with other similar sensors such as SPOT HRV (Systéme Pour l'Observation de la Terra High Resolution Visible) (Lillesand et al. 2008: 399–432; Wulder et al. 2008)

Satellite images along with aerial photographs are the most common sources of raw data for land cover classification, but as such they are only visual representations of the situation at one time and need to be categorized for quantitative analyses (Foody 2004). This can be done manually through digitizing or automatically with computer driven classification techniques. Automated classification techniques are relevantly cost-efficient, therefore suiting regional or national scale studies done with aerial or satellite images. In automated image classification the idea is to categorize pixels to spectral clusters or land cover classes. Classification procedures are mainly based on the differences in DN values in multispectral bands, though some applications also use contextual or texture information in the process (Lillesand et al. 2008: 545–591). Unfortunately even land cover features within same land cover class may have different reflectance properties and for example young birch forest can vary significantly from mature pine forest, although they would both belong to land cover class forest. Therefore it is essential to understand the spectral differences of the features before the actual classification process and to create classification schemes dictating which kind of differences are tried to adduce with the classification. Although even then, some features that are extremely similar in reality may have dissimilar spectral properties because of illumination and shadow differences (Campbell 1996: 316; Di Gregorio 2005: 1). Eventually all land cover classifications include so called omission and commission errors. The first one refers to a situation where a pixel is not classified to that category where it should belong, while the second problem is the opposite where classes contain pixels which are not part of that particular class in reality (Rogan & Miller 2007: 150; Lillesand et al. 2007: 585–586). One approach to these built-in problems of land cover classifications would be to create fuzzy or probability classifications, not indicating the absolute land cover class but a possibility or a probability to belong to a class (Benz et al. 2004).

However just creating a classification does not mean it is ready. Before actually using any classification its accuracy should be tested (Lillesand et al. 2008; 585). This allows users and producers to evaluate maps utility for their purposes (Stehman & Czaplewski 1998). When study areas are large, accuracy is estimated against reference data, which can be collected with multiple techniques from various sources. It can rely on on-site inspections, high resolution satellite images, aerial photographs or other maps from the same region (Lunetta & Lyon 2000; 2–7). Related literature underlines the importance of quality reference data, where spatial resolution is higher than in classified data, temporal gaps are minimal and used classes are recognizable (Lunetta & Lyon 2000; 7). In reality, especially in developing countries the used reference data is more often determined by availability of data rather than rules set in the literature.

3.3. Detecting forest changes from satellite images

Land cover change detection between at least two snapshots of time is the basis for any spatially explicit land change science research. The change detections provide information about the directions, rates, patterns and driving forces of change and although the gained knowledge is primarily historical it can help to detect long-term patterns of changes, which carry influence also in the future (Pontius et al. 2004; Käyhkö & Skånes 2006; Lillesand et al. 2008: 595). Multiple spatial data sources, such as historical and thematic maps, aerial photographs and satellite imaginary can be used in change detection and often understanding long-term patterns dating back decades or even centuries requires combination of all these (Käyhkö & Skånes 2006).

Satellite imaginary change detection is based on overlaying multiple images over each other and comparing them pixel-by-pixel (Coppin et al. 2004; Lillesand et al. 2008: 595). In perfect conditions change detection would be done with images from the same satellite sensor, with the same spatial, spectral and radiometric resolution, taken from the same angle, exactly on the same time of the day and at the same date of a year and eventually registered so that the images overlay each other perfectly. Even if these conditions would be matched, different atmospheric, wind, soil moisture, growing season, rainfall and plant phenology situations between the dates may cause problems (Lillesand et al. 2008: 595). In theory this kind of perfectionism would be possible by using Landsat data (TM or ETM+), but the problems with cloud cover and acquiring of the data from tropics has made it obligatory to rely on data from multiple satellite sensors (Lillesand et al. 2008: 399; Wulder et al. 2008).

Using data from multiple satellite sensors is referred as cross-sensor analysis. Cross-sensor analysis help to produce change detections even when continues single sensor data is not available (Wulder et al. 2008). Though, cross-sensor analyses have their own set of limitations relating to sensor calibration, spectral and spatial resolutions, geo-rectification and temporal coverage (Franke et al. 2006; Wulder et. al 2008). Many sensors like Landsat and SPOT have spectrally rather similar bands, but the calibration differences make direct comparisons of DN values and indices impossible. Comparison of these would require tedious radiometric normalizations, which are often impossible because lack of reference data about the atmospheric conditions or because spatial mismatches and different resolutions makes it impossible to find exactly the same pixels from the two images (Franke et al. 2006; Wulder et al. 2008). All change detection procedures also require geo-rectification of the images. In geo-rectification raw images are registered to known coordinate systems by connecting clearly identifiable elements with Ground Control Points (GCP) from the raw images to same elements in images that are already registered to certain coordinate system. In other words the raw images are overlaid against other spatial data so that they are spatially matching (Lillesand et al. 2008: 485–490; Mather 2005: 88). Generally rectification errors less than ½ of a pixel are considered acceptable, while larger errors may cause dilemmas in change detection if changes are caused by pixel mismatches instead of real world transitions (Lillesand et al. 2008: 595). Also the spatial resolutions vary between different sensors and the images must be resampled to the same resolution, which inevitably causes loss of spatial and substance information and makes the geo-rectification procedure even more difficult (Lillesand et al. 486–488). In other words after any geo-rectification and especially when it is combined with cross-sensor resampling, it is difficult to be secure that pixels laid on each other's represent the same pieces of land, which eventually compromises the whole idea of pixel-by-pixel analyses (Coppin et al. 2004; Lu et al. 2004; Franke et al. 2006; Wulder et al. 2008). For these reason, it is essential that the center of research focus is on large continuous patches and their changes instead of individual pixels (Wulder et al. 2008).

Post-classification comparison is one of the most used methods in deforestation research. The method is able to minimize the differences in atmospheric conditions, sun-angle and sensor calibrations between the used images and therefore do not require laborious radiometric normalization procedures (Lunette & Elvidge 1999: 32; Lu et al. 2004). Though problems related to geo-rectification, spatial resolution and annual differences still exist in post-classification change detection. It has been estimated that the post-classification comparison has a general tendency to underestimate the overall area of change, but where change is detected, it is often overestimated (Lu et al. 2004). Also changes are clear to detect when they happen to certain direction, for example from forests to urban, but when changes are subtle and cause little spectral-radiometric and textural alterations (ea. change from grassland to active field) they may be left undetected (Lambin et al. 2003; Rogan & Miller 2007: 150). In context of forests this means that degradation is often undetectable, even it would not even be so subtle. Olson (1995) was unable to detect canopy coverage decrease of lower than 20–25% in boreal forests, while Souza and Barreto (2000) could not map selective forest harvests in tropical surrounding, even though they knew their location based on field data. The studies were done with Landsat TM and ETM+ data and the problem eventually relates to the pixel size of medium resolution satellite images, it is simply too large to detect disappearance of individual or small group of trees (Healey et al. 2007: 68). Though also other conditions such as ground reflectance, canopy shadowing effect, tree species composition, understory vegetation and topography influenced the errors occurrence (Rogan & Miller 2007: 150–151).

Another aspect relates to the nature of post-classification comparison. The method is able to avoid problems related to direct setting of absolute thresholds for change that has to be done in spectral analyses (ea. 15% decrease in NDVI value would refer to deforestation), but in reality these thresholds are set in creation of classes based on DN values (Coppin et al. 2004; Lu et al. 2004; Lillesand et al. 2008: 550–557, 596–597). For this reason some relevant changes are left unnoticed when they happen close to centers of DN value clusters, but even minor variations may be classified as change if they happen close to the DN cluster boundaries (Lillesand et al. 2008: 550–557, 596–597). In practice this would mean that even extensive forest degradation would be left unnoticed if it would not cross the class border. Detecting subtle changes would require continues or fuzzy land cover classifications based on possibility or probability to belong to a certain category or spectral change detection methods enhanced with other GIS data and high-resolution imaginary (Lambin et al. 2003; Roger & Miller 2007: 150–153). It is also highly time-consuming to create accurate individual classifications and it requires significant amount of reliable reference data.

However the accuracy of post-classification comparison is directly relational to the accuracy of individual classifications (Lu et al. 2004; Coppin et al. 2004). If the overall accuracy is 90% for classification 1 and 80% classification 2 the overall accuracy of the change detection is 72% (90% * 80% = 72%) (Lunette & Elvidge 1999: 32).

Forests can be dynamic, cyclical, linear, secular or reversible in their changes and transition processes are rarely progressive or even gradual (Lambin et al. 2003; Käyhkö & Skånes 2006; Käyhkö et al. 2011). Cyclicity is typical especially to the forests in swidden landscapes where areas are cultivated and then left to forest again (Hett et al. 2012). Using only two snapshots of time is unable to reveal the whole truth of forest dynamics, as it identifies areas simply as deforested, stable or reforested. It identifies land cover situations at two different times, but is often unable to identify the change processes behind these situations. Large trends may become covered by minor short term changes, making explaining, modeling and predicting changes partly biased. Landscape/Land cover Change Trajectory Analysis is a method using multiple snapshots of time instead of only two, thus, able to model also the change processes (Mertens & Lambin 2000; Käyhkö & Skånes 2006, 2008).

3.4. Identification of spatial clusters

Change detection creates maps of the forest change and visual interpretation of these already provide certain idea about its spatial patterns, however these visual interpretations are highly subjective and quantifying the patterns requires other means (Lunetta & Lyon 2000: 7). An easy approach is to simply convert the forest change raster to polygons and to identify the largest polygons, however if the pattern of deforestation is diffusive the processes does not create continuous patches (Geist & Lambin 2001: 69–71; Hett et al. 2012). Multiple spatial statistical methods, which recognize the spatial autocorrelation effect of near-by areas, have been developed. Global clustering methods, such as Moran's I and Getis Ord G, calculate the distances between similar observations and provides one overall figure for the clustering of the whole data set used (Getis & Ord 1992). Moran's I and Getis Ord G vary in such a manner that Moran's I detects overall clustering (value 1), randomness (value 0) and perfect dispersion (value -1), while Getis Ord G detects if it is high or low values that are clustered, but is unable to identify clustering of both simultaneously. Getis Ord G is also unable to detect dispersion or randomness of patterns (Getis & Ord 1992). The global clustering methods only provide overall figures, while local methods, such as Local Indicators of Spatial Association (LISA) and Getis Ord Hot Spot Analyses, also map the clusters. LISA also known as Anselin Local Moran's I calculates Local Moran's I value for each observation. Positive values imply that the observation is part of a cluster, while negative values suggest that it is an outlier among a cluster of other values (Anselin 1995). Getis Ord Hot Spot Analyses looks each value in the context of its neighboring values and compares these values to the average. If the values are significantly higher than average the observation is part of hot spot of high values, while if the values are lower than average is the observation part of hot spot of low values (Getis & Ord 1996).

These methods are good for continous data, but unable to map clusters if there is only observations with the same value (ea. deforestation = 1) (Getis & Ord 1992). Using only single valued observations turns the object from detecting the clusters among multiple values to detecting concentrations of objects from otherwise empty space. Kernel Density Estimation (KDE) is a tool that can be used for this purpose, while it is also able to handle bivariate data (Silvermann 1986). The method calculates the amount of observations within certain cutoff distance and the areas with high amount of observations within the distance receive high values. KDE create outcomes with smoothed surfaces and therefore provides certain fuzziness to the estimation (Silvermann 1986). KDEs have been used successfully in many land use and biology applications such as mapping changes in prehistorical land use, delineating animal home ranges and defining the key areas for conservation (Seaman & Powell 1996; Brinkmann et al. 2011; Grove 2011).

3.5. Linking environmental factors to forest changes

Linking measured deforestation somehow to its causes is seen essential for enhancing the understanding of the drivers of deforestation. It also helps to create better tools for reducing deforestation and for spatial allocation of the resources (Veldkamp & Lambin 2001; Verburg et al. 2006: 117–118). However there are considerable methodological problems in this process. From the 152 deforestation studies meta-analyzed by Geist & Lambin (2001: 18) 76% created these links qualitatively through secondary data and documents published from the study areas, while only 24% did this quantitatively with household surveys or linking secondary data to the process with correlation or regression analyses. Qualitative analyses usually tend to create descriptions or narratives of change process, whereas quantitative analyses rely on structural modeling (Verburg et al. 2006: 117–119).

Multivariate regression modeling have been the most common tool of quantitatively explaining links between deforestation and factors behind it (Geist & Lambin 2001: 18; Veldkamp & Lambin 2001; Verburg et al. 2004b). These methods are able to combine explanatory power of various variables, show the direction of relationships between the dependent and independent variables, determine the power of each individual explanatory variable and allow creating future scenarios based on the regression equations (Veldkamp & Lambin 2001; Verburg et al. 2002, 2006: 119; Walker 2004). The regression analyses can be either spatial or non-spatial. The non-spatial models try to determine the rate and magnitude of change without taking account to its spatial distribution. Spatial models on the other hand try to do this at the level of pixels or administrative units (Verburg et al. 2006: 118). Though even the spatial models have certain spatial differences and for example the models based on administrative units can be purely statistical using generalized deforestation figures as dependent variable and aggregated socio-economic and environmental data as independent variables (Tole; 2001; Aguiar 2007), whereas some GIS –based models explain spatially accurate deforestation distributions with spatially accurate accessibility, distance, soil, elevation, slope and vegetation data (Ludeke 1990; Chomitz & Gray 1996; Geoghegan et al. 2001; Verburg et al. 2004, 2006: 118).

Logistic regressions models are generally the most used regression methods, however they face serious dilemmas when handling spatial data (Kaimowitz & Angelsen 1998: 40–41; Anselin 2002; Verburg et al. 2002: 126–128). Geographical phenomenons are spatially autocorrelated, which may cause problems when determining the importance of the variables or their coefficients. The issue is especially problematic when variables are ranked based on their influence and in some studies independent variables have lost their explanatory value completely after the spatial autocorrelation has been normalized (Rosero-Bixby & Balloni 1996; Kaimowitz & Angelsen 1998: 41). On the other hand spatial autocorrelation is not just a bias, but a real characteristic of geographical phenomenons and therefore not removing it may improve the predictive models (Kaimowitz & Angelsen 1998: 41; Verburg et al. 2002: 127). Tools for both, removing (Spatial Lag and Spatial Error regression models) and incorporating the positive sides (Cellular Automata) of spatial autocorrelation have been developed (Anselin 2002; Verburg et al. 2002: 127).

Also the outcomes of regression analyses might be systematically flawed if there is serious multicollinearity between the independent variables. In other words if there is serious correlation between the explanatory variables this might influence the outcomes (Kok & Veldkamp 2001; Serneels et. al 2007). Some regression methods include automated testing of multicollinearity, however this is not included into the binary logistic regression and therefore needs to be tested before modeling. Correlating variables should be dropped from the final model or combined as one (Serneels et. al 2007). Also the scale influences, not only the binary logistic, but all regression methods. Changes in land cover are results of processes at different scales, issues mattering at certain scale may not explain changes at another scale and aggregating process from down to up do not lead to correct outcomes (Turner et. a. 1995; Veldkamp & Lambin 2001; Verburg et al. 2002: 123–126). Therefore it is a necessity to understand the scale of the study, choose right variables based on it and not to do generalizations that go over the used scale.

Regression analyses are not the only tools for linking environmental factors to land changes, though they have been extremely popular in recent years (Verburg et al. 2006: 116). One different approach is to divide the study area to sub-areas based on elevation, accessibility or vegetation type and to study the differences between the sub-areas (Serneels & Lambin 2001; Nagendra et al. 2003; Southworth et al. 2004). Another approach is to search for more nonlinear connections, which is impossible with the linear regression models (Mertens & Lambin 1997; Verburg et al. 2004a). Often the regression modeling is done for regions where the scarce population concentrates to clearly delineating cities and villages, road networks are limited and the area is largely covered by natural vegetation (Ludeke et al. 1990; Chomitz & Gray 1996; Verburg et al. 2002; Nagendra et al. 2003; Aguiar 2007). These are the perfect conditions for finding linear correlations and to create regression models based on them. However the relations can be far from linear in small and highly populated islands with diverse land use actions. Therefore it was seen necessary to seek ways to study the relationships of deforestation and distance measurements nonlinearly. Related literature did not offer many methodological examples, though some studies that link distance measurements to deforestation in spatially detailed accuracy and in nonlinear manner were found. In their study from southern Cameroon Mertens & Lambin (1997) scatter plotted logarithmic frequency of all the deforested pixels against distances to roads, towns and forest edges and calculated linear, logarithmic and quadratic functions explaining these patterns and used these functions to predict fore coming deforestation. Verburg et al. (2004) linked cumulative distribution of corn, banana and rice farming and secondary and primary forests to travel times to markets at 5-minutes intervals. Maeda (2011: 34–42) worked in slightly different manner and reflected the probabilities for agricultural expansion calculated from regression analysis against individual explanatory variables for better creditability of his modeling.

The spatially explicit regression models along with annual forest cover change rates can be also used as tools for predicting the future of changes (Veldkamp & Lambin 2001; Verburg et al. 2002; Walker 2004). Predictive models provide cartographic information to forest officers and land use planners about areas with highest change potential. These projections can be used to discuss possible changes with relevant stakeholders and eventually to allocate management actions (Mertens & Lambin 1997; Verburg & Veldkamp 2004). The purpose of predictive modeling varies according to the scale. Often at national level it is enough to recognize the hot spots of possible change, while regionally changes should be connected more tightly to overall land use and their influence on natural resources (Verburg et al. 2002; Verburg & Veldkamp 2004). The projections can be either static as certain "business as usual" -models assuming that the developments keep their current track or dynamic estimating certain changes happening in the process (Mertens & Lambin 1997; Verburg et al. 2006: 119; Ahrends et al. 2011). Especially the dynamic models work as virtual laboratories where alternative pathways of future can be tested (Veldkamp & Lambin 2001).

4. Study area

4.1. General geography of Unguja

Zanzibar is a semi-autonomous archipelago in Tanzania and Unguja is the larger of its two main islands (Figure 7). The island is located in East Africa, 40 km east from Tanzanian mainland and slightly south from the Equator (5° - 6° S and 39° E). At longest Unguja is 85 kilometers long and 39 kilometer wide, with total area of 1 660 km^2, making it slightly larger than island of Gran Canaria.

Although Unguja is relatively moist year around, its climate is dominated by tropical monsoon system with two distinct rain seasons. The annual rainfall is around 1000–2500 mm and majority of this is contributed by *masika,* the long rain season from March till June. *Vuli,* the short rains, lasts from October until December and provides approximately one third of the rainfall. Eastern side of the island gets relatively less rainfall and *Vuli* can be completely absent in some years (Hettige 1990: 11; Krain 1998; Klein & Käyhkö 2008). The annual mean maximum temperature is 29,3°C and mean minimum 21,1°C. The hottest period is from January until February in the dry season just before the long rains, while the coolest period is between May and September, from the end of the long rains until the beginning of short rains (Hettige 1990: 24; Krain 1998; Klein & Käyhkö 2008). Geologically the islands can be divided to two larger regions: to the elevated and undulating terrain with fertile deep sandy soils and maximum elevation of 120 meters in the western side and to the generally flat, coralline limestone dominated area with terrace system descending stepwise from 40 meters until the sea level in the eastern side (Hettige 1990: 25–30; Klein 2008a: 15). Though the eastern side is generally coralline limestone dominated, the composition of soils vary rather randomly from hard and solid coral rag with only pockets of softer soils to mixture of broken down coral rag and more fertile soils (Klein 2008)

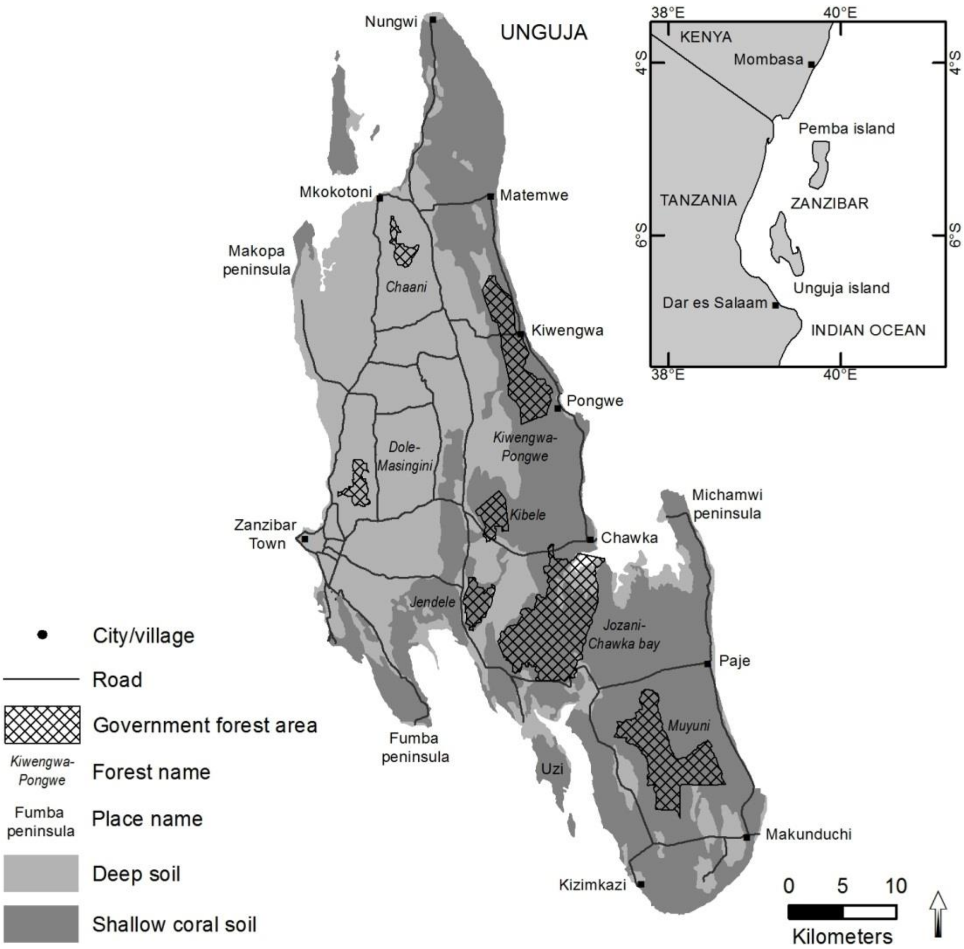


Figure 7. Island of Unguja in Zanzibar, Tanzania. The island is roughly divided to deep soils in the west and shallow coral soils in the east. There are all together six government forest conservation or plantation areas.

In the population census of 2002 there were 980 000 habitants in Zanzibar and 620 000 in Unguja. The average population density of 382 people/km^2 makes the island one of the most densely populated rural areas in the world, and the annual growth rate of 3,1% guarantees increase of population for years to come (OCGS 2007: 1–10). Based on the annual growth rate between 1988 and 2002, the population in Unguja today should be close to 850 000 and the average population density should be over 500 people/km^2. Zanzibar's' GDP was as low as 543 USD in 2008, although the real gross domestic product (GDP) has grown impressive 6,4% annually since 1991 (RGZ 2009: 29–31; OCGS 2010: 3–4). About half of the households are below basic needs poverty line and 13% of them fail to sustain daily food consumption. However, the situation is not as bad as the poverty figures may imply, since subsistence uses of land, good access to water, healthcare and education reliefs the situation (RGZ 2006).

The economy of Zanzibar has been liberalized since 1990s and the annual GDP per capita growth has been astonishing 6,4% between 1991 and 2008 (RGZ 2009). The GDP is divided between services (42,7%), primary production (30,7%) and industry

(14,3%). However GDP does not include subsistence production and thus underestimates the importance of primary production. Primary production comes mainly from crops (21,3%), livestock (4,6%) and fishing (4,4%), while forestry production (0,3%) has only a marginal share (OCGS 2010: 5). Clove, seaweed and rubber were the main cash crops, while cassava, banana, sweet potato, coconut and rice were grown for food consumption. Forestry generated incomes mainly through selling of wood fuel and charcoal, while building poles and other forest materials created only ¼ of the income (OCGS 2010: 18–20). The most important sectors outside primary production were public administration (9,7%), trade and repairs (8,7%), transport and communication (8,0%), hotels and restaurants (7,4%), and construction (7,3%) (OCGS 2010: 5).

It is not the history as center of slave trade, period as the “Spice Islands” or modern poverty that first pops up to people’s minds about Zanzibar, but rather the status as tourist destination. Tourism has grown massively since mid-1990s, especially in the east coast of Unguja (Mustelin 2008; Gössling 2001). The spread of international hotel chains have not only increased the price of land in the coast, but also restricted access to communal land, limited possibilities to fishing and sea weed farming and caused direct and indirect deforestation. Though, Zanzibar generates approximately 25% of its revenues from tourism, it does not create equal employment effect and the benefits to the coastal communities have been rather limited (Mustelin 2008; RGZ 2009: 17)

The land tenure system in Zanzibar is rather confusing. In theory all the land is nationalized by the government and only rights to land use are given to habitants. These rights can be acquired from village leader and are ratified via actual use of land (ea. building a house). After land use rights are acquired the holder can sell or leave the constructions, trees or other improvements as an inheritance. This structure is mixed up by the government land redistribution programs and the inheritance traditions of Islam (Krain 1998; Törhönen 1998; Fagerholm 2012: 34–35). All together the baffling tenure systems, causes unwillingness to long-term land management, creating a perfect starting point for the “tragedy of commons” (DCCFF 2008).

Behind generalizations of an entity there are significant differences within Unguja. Administratively Unguja is divided to 3 regions, 6 districts and 197 wards (*Swahili: shehia)* (Figure 8). One-third of the population are concentrated within the small Urban district covering central parts of Zanzibar Town, another one-third within West district and final one-third are gathered between the four other districts, which occupy 86% of the land area (Table 4) (RGZ 2007: 23; RGZ 2009: 15–16). Highly concentrated population creates significant differences in the population distribution and the population density for Urban district is over 146 times higher than it is in the least populated South district. The differences are significant also within the districts and majority of buildings are concentrated along the coast or the main roads (Figure 8). District specific population growth rates varies between 1,7% (South) and 9,2% (West). High figure of West district and relatively low figure for Urban district (1,9%) indicates that Urban district is not anymore capable to accommodate more inhabitants and majority of rural-urban migration and urban sprawl happens within the West district (Ameyibor et al. 2003: 7; NBS 2005).

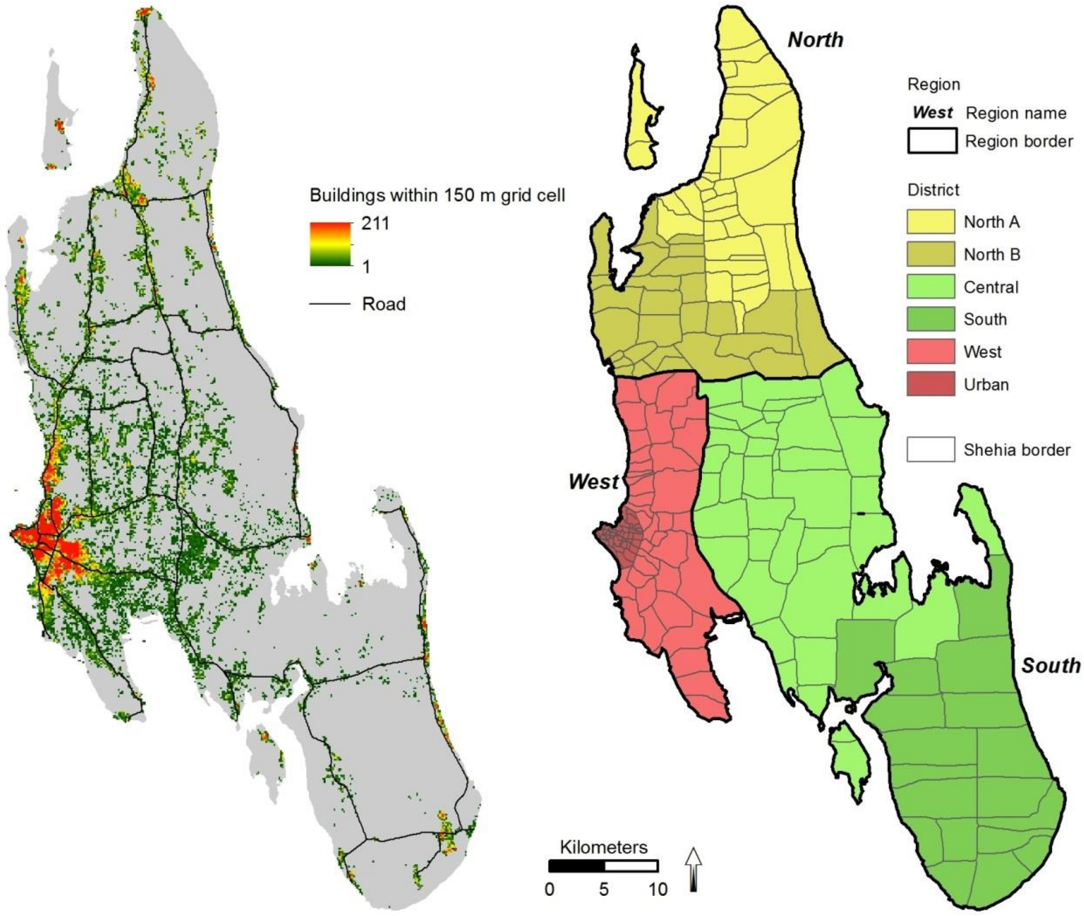


Figure 8. A) Distribution of buildings within 150 m² grid cells in 2004 calculated from DoSUP 2009 building database. B) Administrative borders of Unguja. Data source: DoSUP 2009 topographic map database.

Table 4. Demographic, socio-economic and agricultural statistics of the five districts of Unguja (Based on NBS 2004a, 2004b, 2004c, 2004d, 2004e & 2004f)

Population	**North A**	**North B**	**Central**	**South**	**West**	**Urban**
Total population	84 147	52 492	62 391	31 853	184 204	205 870
Urban population (%)	1,1	2,7	1,5	12,2	61,7	100
Population density	399	244	138	88	886	12 867
Population growth (%)	2,4	2,4	2,3	1,7	9,2	1,9
Employment						
Primary production (%)	48,8	39,9	43,6	30,3	11,4	3,2
Government employed (%)	3,5	7,0	6,2	6,3	12,3	13,1
Private sector employed	2,5	2,1	2,2	3,4	8,3	8,2
Unemployment rate (%)	3,4	5,5	1,7	2,3	8,4	14,7
Dependency ratio	102	89	88	85	83	68
Economy						
Annual mean per capita income (Tsh)	159 786	177 578	159 226	181 942	235 548	271 915
Agricultural share of income (%)	14,3	17,3	24,9	11,2	6,2	1,8
Monthly consumption per capita (Tsh)	18 099	16 667	19 901	18 134	23 105	28 745
Consumption on food (%)	60,4	60,7	62	62,5	54,6	50,9
Population below food poverty line	12,18	12,06	8,35	9,73	9,54	7,75
Population below poverty line	53,3	48,3	45,7	53,8	38,6	37,6
Education						
Adult literacy (%)	46	63	71	75	77	81
Adults with no education (%)	43	33,3	19,6	16,2	9,3	12,1
Healthcare						
Infant mortality rate	113	87	85	92	70	69
Child mortality rate	84	57,5	56	62,5	42	41,5
Household						
Average household size	5,3	5	5,2	4,7	6,1	5,4
HHs with modern roof (%)	60,8	38,2	53,5	59,2	74,2	92,7
HHs with modern walls (%)	56,5	39	19,2	11,9	76,2	73,5
HHs with electricity (%)	4,1	7,9	6,5	19,5	34,1	67,6
HHs without toilet facilities (%)	57,8	40,6	24	29,4	6	0,7
Use of firewood in cooking (%)	97,4	96,01	96,85	92,1	57,32	42,56
Primary production						
Average cultivation area per HH (ha)	0,8	0,85	0,92	0,44	0,6	n/a
Agricultural households (%)	88,9	83,6	92,9	62,5	34,9	n/a
Households growing crop (%)	88,7	83,4	92,9	61,9	34,4	n/a
Households rearing livestock (%)	18,8	28,5	38,8	18,4	13,9	n/a
Annual crop per person (tons)	0,13	0,14	0,16	0,06	0,03	n/a
Fish catch per person (tons)	0,044	0,004	0,016	0,040	0,006	0,028

Based on the district level statistics Urban and West districts can be generally characterized as urban areas with higher income, higher share of urban livelihoods, better education levels, better healthcare, more modern housing, but also higher unemployment ratio, higher population growth and higher percentage of young people. The West region has Human Development Index value of 0,86, while in the other two regions it is below 0,65 (RGZ 2009: 16). Urban and West are more developed districts, which are also facing problems related to urbanization. Majority of agricultural and forest products are brought from other areas and direct dependency on primary production is lower, though gradual change towards rustic can already be seen in the West district.

The four other districts can be characterized as more rural and primary production based with less population pressure and lower development status. The lower incomes are compensated with domestic food and natural resource production, but still poverty and incapability are more present here than in urban areas. Fishing is an important source of livelihoods in North A, South and Urban districts, while the people in South are less engaged to agricultural actions. If literacy, education and food poverty are used as standards, Central and South districts are doing better than North A and North B and especially North A appeared as the least developed area in the light of urbanization, share of agricultural production, income, food poverty and healthcare.

4.2. Landscapes of Unguja

The landscapes of Unguja can be roughly divided to two types: the western deep soil and the eastern coral rag landscapes (Figure 9). The coral rag region is characterized by cliffs, terraces and less fertile coral soils, which are limited to support permanent agriculture, causing the landscapes to be dominated by shifting cultivation and natural vegetation (Hettige 1990: 95–98; Kombo & Kitwana 1997; Klein 1998; Klein & Käyhkö 2008). Majority of tree species are indigenous, but because of the extensive shifting cultivation, fire wood collection and other land use pressures the trees rarely have time to grow to mature forests. Thus the landscape is highly fragmented and reflects different stages of succession varying from recently used fields, fallows, scrubs, thickets to various forms of forest. The heterogeneity of the landscape is enhanced by distribution of settlements, the permanent agriculture and agroforestry around them and the uneven distribution of fertile soils (Krain 1998; Klein 2008b; Käyhkö et al. 2011; Fagerholm 2012: 37). Old-growth forests are mainly present at sacred sites and formally protected areas, the vastest being government forest reserves of Kiwengwa-Pongwe and Jozani (RGS 2004).

Cassava, corn, jams, tomato, and beans are the most common agricultural products used in shifting cultivation, but also fruit trees are planted. The shifting cultivation rotation is rapid (1 – 5 years) and it is not supported with machines or fertilizers, making it extensive and covering large areas (Fagerholm & Käyhkö 2009). Besides agriculture, the land is used for collection of fire wood, building poles, medical plants, handicraft materials, charcoal, coral stone, lime, wild fruits and vegetables, grazing of livestock and beekeeping. The end products of these activities are occasionally sold for income (Orjala 2006: 30–31; Fagerholm & Käyhkö 2009; Fagerholm et al. 2012). Use pressure concentrates to the vicinity of settlements, which are subjected to degradation, deforestation and fragmentation of landscapes (Fagerholm et al. 2012). Almost all villages are in coastal settings, therefore fishing, sea weed cultivation and other marine based livelihoods are also important (Orjala 2008; Fagerholm 2012: 37). Besides the traditional livelihoods, tourism development has occurred mainly at the Eastern coast of Unguja (Gössling 2001; Mustelin 2008).

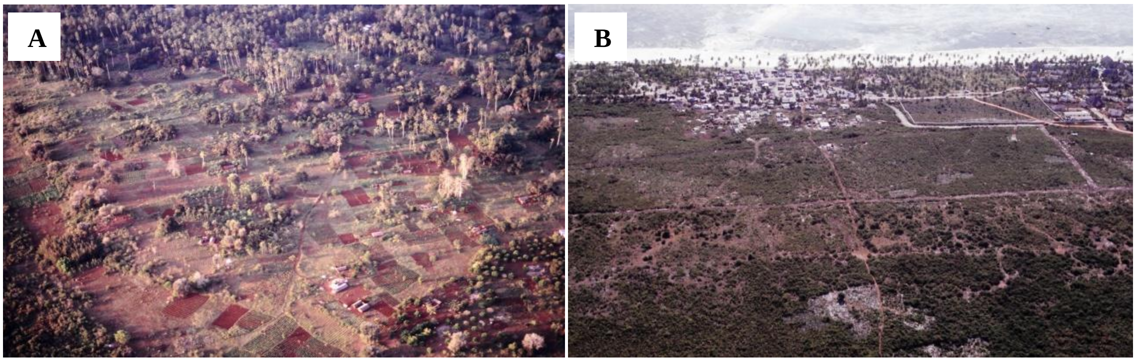


Figure 9. A) Deep soil agroforest system where trees are mixed with open cultivation. B) Coral rag coastline with scrubs of various heights and structures. Photos: Jukka Käyhkö 2004.

The western part of the island is characterized by elevated and undulating terrain with fertile deep sandy soils, making it the base of permanent agricultural production of Unguja (Klein & Käyhkö 2008; Klein 2008). Although the superiority of the western deep soils has been questioned, majority of the agriculture and population concentrates there and makes natural forests rare (Klein & Käyhkö 2008). It has been assumed that the region consisted mainly of tropical forests before it was turned into food crop agriculture and plantations coconuts and cloves in the 19th century (Hettige 1990; WWF Tanzania Country Office 2012: 59). Today the landscape is a mosaic of fruit tree plantations, agroforestry, permanent cultivation, settlements and other infrastructure, while natural forest can be only found from the government protected area of Dole-Masingini (Hettige 1990: 85–94; Klein & Käyhkö 2008). Mainly the area is covered by semi-open fruit plantations and agroforest systems except for the open cultivation and urban areas (Hettige 1990: 85–94). Krain (1998) distinguished the hydromorphic valleys situated in the middle of the island dominated by irrigated rice cultivation, as a third agro-ecological/landscape zone. In these valleys the agroforestry system turns to open cultivation and tree planting is actually forbidden. Besides these valleys, one could argue that Zanzibar Town has expanded to such an extent that it could called as its own landscape zone within the Deep soil region.

In the deep soil region the main agricultural crops are rice, cassava, maize, sweet potatoes, banana, cowpeas, pigeon peas, pineapples and various vegetables, while mango, coconut, papaya, cloves, durian, kapok, breadfruit, jack-fruit and citrus trees are grown in agroforestry and fruit plantation systems (Hettige 1990: 85–94; Orjala 2006: 30). Significant portion of the agricultural production is sold to Zanzibar Town or in the local markets and altogether the vicinity of the capital allows more diverse livelihood options than in more distant coral rag region. Otherwise the rural livelihood options are rather similar except fishery and forestry based livelihoods are less common in the deep soil area (Orjala 2006: 30).

4.3. Forests of Unguja

Unguja's forests are part of the Coastal Forests of Eastern Africa, which are classified as global biodiversity hot spot, because of their high diversity of endemic plants and animal species, such as Zanzibar Red Colobus Monkey (*Procolobus kirkii)* (Burgess & Clarke 2000: 71–73; Siex et al. 2011; WWF Tanzania Country Office 2012: 21). However, on itself Zanzibar Islands are rather depauperate in the sense of endemic species, mainly because of at least of 2000 years of human influence. There are only four endemic and 93 regionally endemic species in Unguja (Burgess & Clarke 2000: 137–142). It is assumed that without any human interventions the island would be covered by tropical high forest in the deep soils and deciduous woodland in the coral rag region (Hettige 1990). Today the forests are mixture of evergreen and deciduous trees and scrubs, woodlands and various thickets (Figure 10) (Burgess & Clarke 2000: 84–94). Estimations of forest cover for whole Zanzibar vary between 660 km^2 (26%) and 1350 km^2 (51%) depending on how forest is defined (RGZ 2004: 6; DCCFF 2008: 39). In the Biomass Inventory of 1997 the forest cover was divided between Unguja and Pemba so that the former had 69% (928 km^2) of the terrestrial natural forest cover if agroforest and mangroves were left out from the calculations. Based on this percentage cover of natural forest would be 455 km^2 (27% of total landscape) in the DCCFF estimation from 2008. The biomass inventory also divides the forests to coral rag forest (65%), agroforests (30%) and high forests/plantations (5%) (RGZ 2004: 6).

Although forestry was only 0,3% of the total GDP it has an important role in subsistence consumption and such materials as fuel wood, charcoal, wood for lime making, building materials, handicraft materials, medicinal plants and wild fruits are collected from the forests (Orjala 2008; Fagerholm & Käyhkö 2009; Fagerholm et al. 2012). The forests also have significant aesthetic, cultural, spiritual, religious and intrinsic values to the local communities and especially the Masingini Forest reserve has an important role in securing the quality of the ground water used in Zanzibar Town (Fagerholm et al. 2012).

Figure 10. Forest area in the outskirts of Zanzibar Town steadily becoming younger and shorter because of extensive use.

When talking about the forest of Unguja the focus is often placed on natural forests, while domestic forests are left with less attention. However, approximately one-third of the forests are within agroforest systems, consisting of coconut, mango, clove, rubber, orange, durian and rambutan trees along with the indigenous species (RGZ 2004: 6; OCGS 2006: 172–174). In small scale agriculture coconut is the most common tree planted, while also orange, mango and clove are also widely grown (Table 5). Together the area used for these crops cover 4,8% of the total landscape, while regional differences are existing. Proportionally coconut is the prominent tree crop in Central district, orange in the South, mango in North A and clove in North B, while planting of agroforest or fruit trees is generally more rare in the South district (1,1% of total area). Also government and large scale agricultural producers are involved with rubber, coconut and clove production, but statistics are not available. Besides agroforestry, 5% of small scale agricultural households are engaged in their own tree planting activities and 3% in communal tree planting schemes. The most often planted trees are Casuarina *(Casuarina equisetifolia)* and Acacia *(Acacia auriculiformis)*. Majority of communal and individual tree planting is done in North A and Central districts, and especially communal tree planting schemes are rare elsewhere. Households own tree planting is done mainly for producing poles and planks, while fuel wood is seen as important secondary product. Communal tree planting is rationalized as well with material reasons, but in Central and West district it has also environmental rehabilitation and erosion control purposes (OCGS 2006: 218–220).

Table 5. Agroforest tree crops in different districts based on OCGS 2006: 172–174.

	Area (ha)					**Percentage of total landscape**				
District	Coconut	Orange	Mango	Clove	Total	Coconut	Orange	Mango	Clove	Total
North A	514,6	129,0	653,2	89,4	1386,2	2,4	0,6	3,1	0,4	6,6
North B	529,9	45,6	68,8	368,6	1012,9	2,5	0,2	0,3	1,7	4,7
Central	1251,3	1119,2	460,2	220,7	3051,4	2,8	2,5	1,0	0,5	6,7
West	865,2	248,9	57,9	62,8	1234,8	4,2	1,2	0,3	0,3	5,9
South	181,4	129,8	73,8	0,0	385,0	0,5	0,4	0,2	0,0	1,1
Total	3342,4	1672,5	1313,9	741,5	7070,3	2,3	1,2	0,9	0,5	4,8

DCCFF (2008) estimated annual rate of forest cover decline to be -1,2%. Also forest degradation and fragmentation are seen as serious threads, but absolute estimations have not been made (Siex et al. 2011). Agricultural expansion, mainly in form of shifting cultivation, collection of wood for energy (fuel wood & charcoal), lime making, building materials and expansion of urban structure and tourism infrastructure are seen as the main direct causes behind deforestation (RGZ 2004: 6, Käyhkö et al. 2008: 73–74; Siex et al. 2011; WWF Tanzania Country Office 2012: 69–70). All these driving forces have their own distinctive spatial behavior. Tourism causes deforestation directly near the coastline, but simultaneously pushes local inhabitants to resettle more inlands, causing indirect deforestation there (Mustelin 2008; Käyhkö et al. 2011). Zanzibar Town grows along its main roads causing direct deforestation mainly at its outskirts in West district, but the growing demands of urban population indirectly increases the forest pressure far away from the city (Masoud 1991; Ameyibor et al. 2003; Ahrends 2010). Fuel wood collection on the other hand happens close to the rural villages or there were wood is easily available (Fagerholm & Käyhkö 2009; Fagerholm 2012). The underlying causes of deforestation are not profoundly studied, but population growth, insecure land tenure system, low technological development, weakness of government institutions and limited livelihood and income generating opportunities are often mentioned as reasons behind unsustainable forest use (Käyhkö et al. 2011; Siex et al. 2011; WWF Tanzania Country Office 2012: 69–70).

Department of Forestry and Non-Renewable Natural Resources (DFNR) is the government official responsible for forest resources in Zanzibar. Its main goals are to protect biodiversity of forest and mangrove areas, secure sustainable use of forest resources, develop farm forestry and enhance the capacity of forest management. The coral rag forests are set under special attention, because past difficulties securing their sustainability (DCCFF 2008; Fagerholm 2012: 34–36). In the past DFNR has been able to implement large scale tree planting schemes and gazetting of protection areas, while focus nowadays is more in management of the already established forest reserves, increasing their connectivity and ensuring sustainability of communal forests (ZFDP 1997; DCCFF 2008; Siex 2011).

Based on by the official delineations of government forest areas provided by DFNR 103,1 km^2 of land is under official government protection and 20,5 km^2 is used for government forest plantations. There are three established sites of government protection: Jozani-Chwaka Bay National Park (66,9 km^2), Kiwengwa-Pongwe Forest Reserve (30,4 km^2), Masingini Forest Reserve (5,8 km^2) and one newly proposed forest reserve of Muyuni (42,1 km^2) (Siex 2011; WWF Tanzania Country Office 2012: 59). All the important forest areas, expect for Masingini, are connected by corridors in community lands, but deforestation is risking these connection if major actions are not taken (Siex et al. 2011). There are also other Government Forest Areas, such as Chaani (4,5 km^2), Dunga (7,9 km^2) and Kibele (8,1 km^2), which are tree plantations not meant for conservation. Possibilities to manage and control forest uses are limited in governmental areas, due to lack of monetary resources and the situation is even worse in communal lands. In some areas, mainly Jozani-Chwaka Bay National Park, revenues created through ecotourism are helping the situation (WWF Tanzania Country Office (2012: 70).

At least half of the biologically important forest habitats lay outside the government areas and the communities are brought into management and conservation of forests (Siex 2001; WWF Tanzania Country Office 2012: 69–70). The process started already in 1990s and at the moment it is implemented through Community Forest Management Agreements (CoFMA), which zones forest areas and gives legal management mandate to the communities. Simultaneously it requires communities to set the responsibilities of protection and sustainable use, which should be then followed (WWF Tanzania Country Office 2012: 85). The CoFMAs restrict open access to forest land by dividing it to high protecting, low impact and high impact zones and thus providing community based solutions for “the tragedy of commons” (Hardin 1968). Though funding to fulfill the laid responsibilities is still uncertain and implementation has been criticized because of limited public participation (Fagerholm 2012: 36). CoFMA process is also supporting global REDD+ actions, where Zanzibar is piloting as a part of Tanzania (The Royal Norwegian Embassy 2010).

5. Materials and methods:

5.1. Primary data

Two satellite images were used as primary source for the land cover classifications, SPOT-3 HRV image dated 30.06.1996 and Landsat-5 TM image from 01.07.2009 (Figure 11). The Landsat image was already obtained for previous research and georectified according to the 2004-2005 aerial images, while the SPOT image was purchased from commercial company (Astrium) for this purpose, since other cloud free Landsat data was not available. The SPOT image was acquired from the same time of the year than the Landsat TM image to minimize the seasonal variations in vegetation and reflectance. Landsat TM image has six spectral bands with spatial resolution of 30 meters and the thermal infrared band in 120 m resolution, while SPOT image has only three bands all in higher 20 m resolution (Table 6). The wavelengths missing from the SPOT image in relation to the TM image are blue, two mid-infrared and the thermal-infrared bands. Arc 1960 UTM Zone 37S was used as the coordinate system for both satellite images and also for other spatial data from Zanzibar. The created classifications were later on used for change detection.

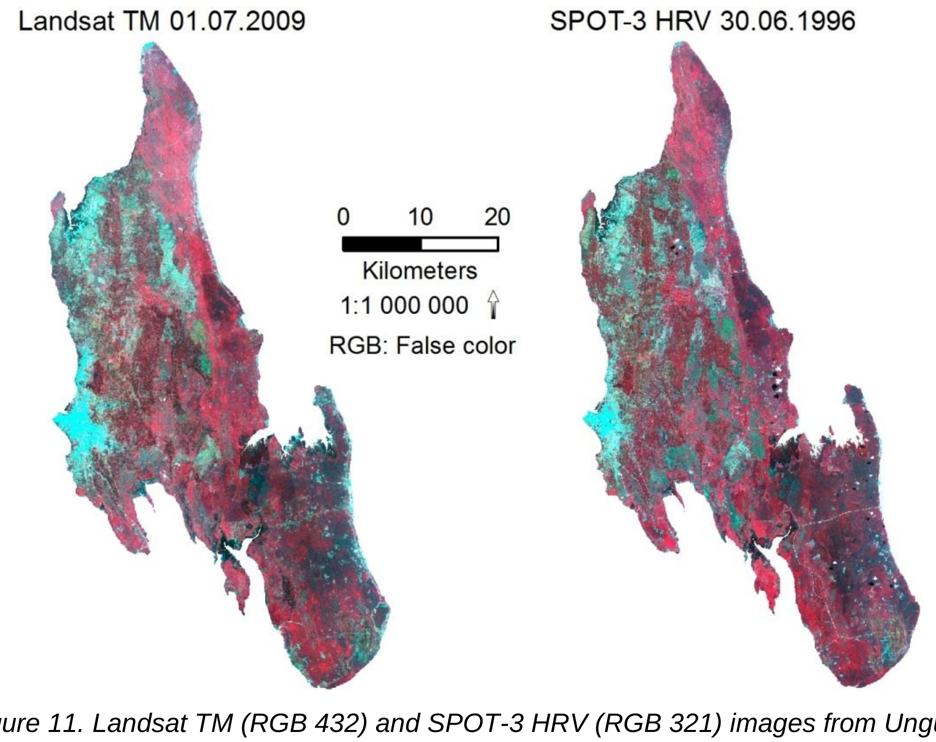


Figure 11. Landsat TM (RGB 432) and SPOT-3 HRV (RGB 321) images from Unguja.

Table 6. Spectral reflectance and spatial resolution of Landsat TM and SPOT-3 HRV satellite sensors. Based on Lillesand et al. 2008; 411, 433.

Band	Wavelength (µm)	Spectral location	Resolution (m)
		Landsat TM	
1	0,45 - 0,52	Blue	30
2	0,52 - 0,60	Green	30
3	0,63 - 0,69	Red	30
4	0,76 - 0,90	Near-infrared	30
5	1,55 - 1,75	Mid-infrared	30
6	10,4 - 12,5	Thermal-infared	120
7	2,08 - 2,35	Mid-infrared	30
		SPOT-3	
1	0,50 - 0,59	Green	20
2	0,61 - 0,68	Red	20
3	0,78 - 0,89	Near-infrared	20

5.2. Reference data

Aerial photographs of years 1989 - 1990 and 2004 - 2005 from the Department of Urban and Rural Planning (DoURP) (formerly Department of Survey and Urban Planning) and high resolution GeoEye-1 satellite images received from GeoEye Foundation from year 2009 were used as reference data to assess the accuracy of created classifications (Figure 12). The aerial photographs were already orthorectified, while the GeoEye-1 satellite images were georectified to the aerial images. The 2004 - 2005 aerial photographs were color images with 0,5 meters spatial resolution covering the whole island (Table 7). The 1989 - 1990 images were panchromatic with 1 meter resolution data covering only selected sites of the study area. The 2009 GeoEye-1 high resolution satellite images had five bands, 3 bands of visible light, near-infrared and panchromatic with different resolutions. These images were obtained only for Kiwengwa-Pongwe Forest Reserve and Muyuni-Uzi area, mainly because these were the only cloud-free images close to the date of the Landsat TM image.

Figure 12. Examples of the aerial photographs and high resolution satellite images used in the accuracy assessment. A) Panchromatic aerial photograph from 1989-1990. B) True-color aerial photograph from 2004-2005 C) Multispectral high resolution satellite imaginary from 2009.

Table 7. Spectral band, resolution and spatial coverage of the used aerial photographs and high resolution satellite images.

Spectral band	Resolution (m)	Coverage
2004 - 2005 aerial photographs		
Red	0,5	
Green	0,5	Unguja
Blue	0,5	
1989 – 1990 aerial photographs		
Panchromatic	1	Dole, Cheju, Matemwe, Kiwengwa
2009 GeoEye-1 satellite image		
Blue	1,65	
Green	1,65	
Red	1,65	Kiwengwa-Pongwe (04.10.2009) & Muyuni-Uzi (19.06.2009)
Near-infrared	1,65	
Panchromatic	0,41	

5.3. Field work data

Land cover classification was enhanced with field observations collected between October and December of 2011. Altogether 85 field observation points were visited and premade field observation sheets were filled from each site (Figure 13).

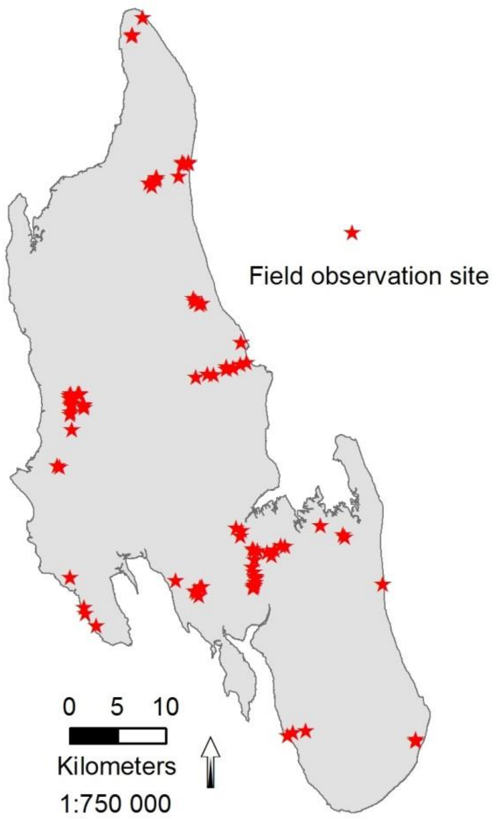


Figure 13. Locations of the collected field observations.

Such issues as topography, soil, moisture conditions, ground layer, number of tree and scrub layers, dominating canopy shape, tree/scrub coverage, spatial pattern, dominant tree/scrub species, ongoing land use activities, evidence of past land use activities and land cover type were visually observed (Appendix 1). Also photos and GPS points were taken from every site. The field work points were used to connect the spectral cluster of unsupervised classification to actual land cover classes and to describe the created land cover classes through statistics, qualitative descriptions and photographs.

5.4. Other spatial data

Also other spatial datasets were used in the analysis process. Multiple thematic GIS datasets from the DoURB 2009 topographical database, such as coastline, roads, building and land use were used (Table 8). All these materials were originally digitized at the scale of 1:10 000 from the 2004–2005 aerial photographs and they were in shapefile form. Coastline shapefile included the accurate borders of Unguja Island as polygon data. The coastline data combined with the mangrove delineation data from DoURB 2009 land use classifications and used to determine the accurate borders of Unguja and to exclude all the island that were not connected to the main island by land or mangroves. Coastline dataset combined with mangroves were also used to create the distance to coastline surface and for enhanced visualization of created maps.

Table 8. Other spatial datasets, their sources and use.

Data	Source	Use
Coastline	DoURB 2009 topographical database	Delineation of Unguja, distance to coastline surface and visualisations of maps
Road lines	DoURB 2009 topographical database	Distance to roads surface and visualisations of maps
Building polygons	DoURB 2009 topographical database	Kernel density of buildings surface, density of buildings map
Land use classification	DoURB 2009 topographical database	Planning of land cover classification, enhance classification of rubber plantations, delineation of Unguja.
Physiographic map	Hettige (1990) & Klein (2008)	Delineation of deep soil and coral rag landscapes
Government forest stations	DFNR (2009)	Delineation of government protected and silviculture forests
Muyuni Forest Reserve	DFNR (2011)	Delineation of planned Muyuni forest reserve

The road database was in polyline form and included all roads visible in the 2004–2005 aerial photographs (24 522 lines) classified to footpaths, tracks, secondary roads, main roads and main roads with two lanes. The database also included length data for each line digitized. Roads data set was used to create the distance to roads surface and in the visualization of some maps. The building dataset included all buildings digitized as polygons divided between Zanzibar Town (76 047) and rest of the island (102 261 polygons). The data set included information about the size and perimeter of polygons, their construction status (build/under construction), block number, block name, section name, class (residential, mosque, church, hospital, school, dispensary, government office, market, police station, etc.) and a name for known buildings. The data set was used to create the kernel density estimation of buildings used in regression analyses and for the building distribution map. The land use data included 17 land use classes: airport, clove, coconut, cultivation, forest, mangrove, mixed, park, plantation, quarry, rice, rubber, salt farm, sand, scrub, settle and wetland. However the classification was created for whole Zanzibar and therefore some of these land uses were missing from Unguja. The dataset was in polygon form and included knowledge about the size and perimeter of each individual patches. The land use dataset was used to support the planning of land cover classification, enhance the classification of rubber plantations, and include mangroves in coastline delineations. Physiographic map of Unguja created by Hettige (1990) and digitized by Klein (2008) in previous research was used to delineate coral rag and deep soil regions. The physiography was divided to four main classes: Alluvial system, marine system, ridge system and shallow coralline system and altogether there were 22 sub-classes. Alluvial system, ridge system and marine system (excluding mangroves) were combined as deep soil system and shallow coralline system represents the coral rag region. The official borders of government borders were provided by the DFNR in shapefile form. The data is collected from the field by following the border marks of the forest areas by foot. DFNR also provided the preliminary border of Muyuni Forest Reserve which follows the forest delineations made in CoFMA process. The datasets include name and area information for each polygon.

5.5. Study approach

The four main research questions are divided to eight sub-questions, which are answered by different methodological tools (Figure 14). The research proceeded in stages and the outcomes of previous work steps were used as the main data for the following ones. The work starts with quite normal remote sensing methods of land cover classification and change detection and their results are post-analyzed with more GIS and geostatistical oriented methods of Kernel Density Estimation, sub-area, distance and regression analysis. The questions related to the future of forests are not answered by any own methods, but rather as byproducts of other methods and research questions.

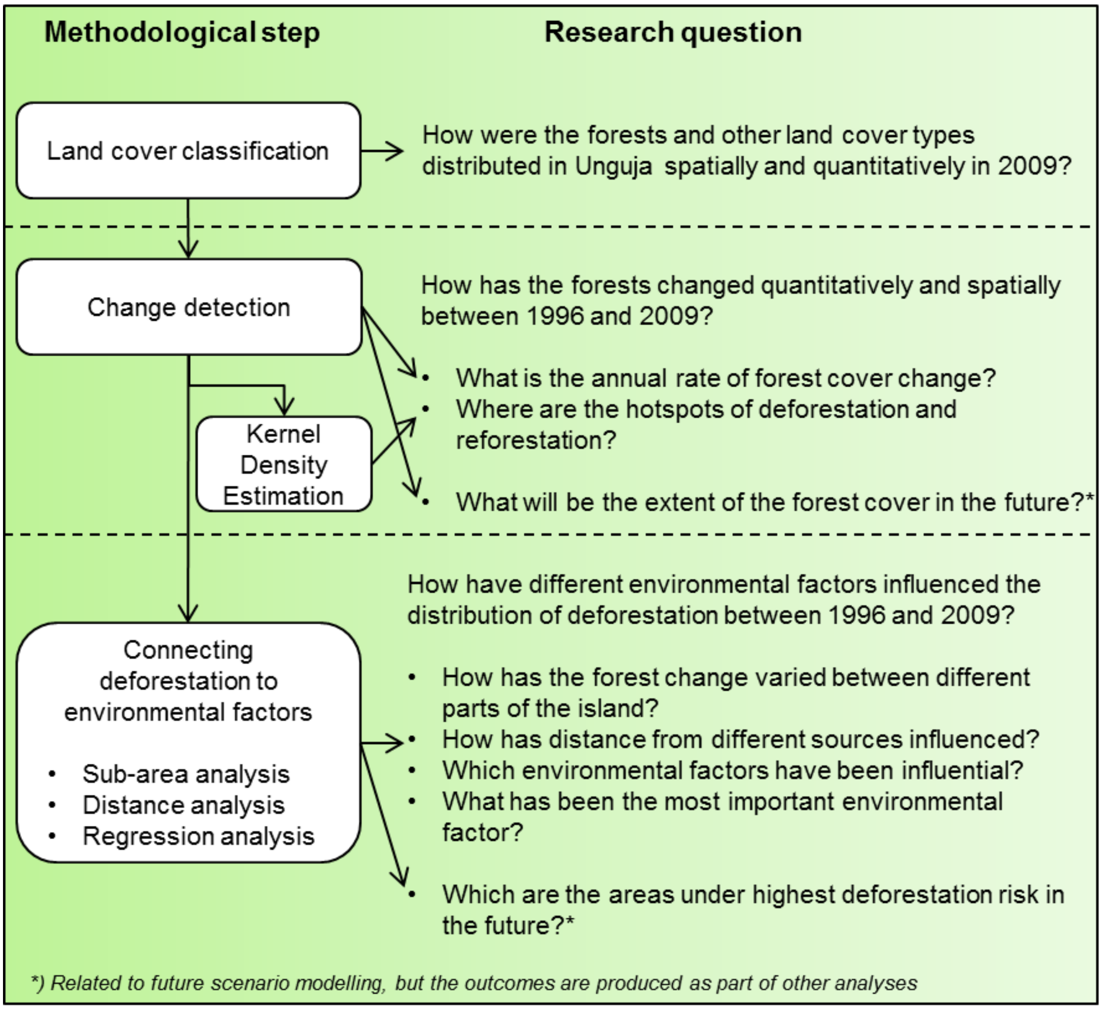


Figure 14. Study approach linking the used methods to research questions.

The work started with the preprocessing of the primary data, which included the acquiring and georectification of the images. This step was followed by land cover classification, which included the planning of the classification scheme, unsupervised clustering, supervised classification, calculating mean NDVI values for classes, accuracy assessment and connecting field observations to classification. Although the classifications included multiple classes the focus was constantly to classify forests as accurately as possible. The two land cover maps created were used in the change detection of forests with post-classification comparison method, which produced change data with 49 possible directions of change. This dataset was used to calculate statistics related to forest cover change, but dataset with 49 classes was considered too complicated and it was developed to two change classifications with fewer categories. These two change classifications were used to create forest change maps, calculate statistics of change, determine the rate of forest cover change and to estimate the extent of forest cover in the future. The change detection outcomes were also used to map the clusters of deforestation and reforestation with Kernel Density Estimation method and to link the environmental factors to deforestation.

Environmental factors were linked to forest change with three different methods: sub-area, distance and regression analyses. The outcomes of regression analyses were also used to estimate the deforestation risk of different areas. Each methodological step answer to specific research questions or their sub-questions and in some cases multiple methods were used to answer same question to give more in-depth perspective.

5.6. Preprocessing of the raw images

Because of lack of other cloud free Landsat images, SPOT-3 HRV image had to be acquired for the change detection purposes. The image was purchased and received in two separate images, where the northern and southern parts of the island were separated. The separated bands and images were stacked and mosaicked to one single file containing all the three spectral bands and both parts of the island in Erdas Imagine 11. The mosaicked image was rectified, registered and resampled to match the Landsat image.

Image was georectified with Erdas Imagine “Control Points” – tool. The tool works in such a manner that the user looks for recognizable points from the raw image and their counterparts from the reference data and places GCPs to these locations. The location in both data sets are recorded and these locations are used in least square regression analysis, which determines regression coefficients that can be used to relate the raw image to the coordinate system of the reference data (Lillesand et al. 486). Altogether 15 GCPs were evenly distributed around the island to locations easily detectable from both images, such as island tips, corners of large build areas and road intersections. Although it is often noted that lower resolution images should be rectified against higher resolution images, the SPOT was rectified against the Landsat TM scene, because it was already rectified to match other available spatial data. The accuracy of the rectification is measured with roots mean square (RMS) error, which is the difference between desired GCP location and the actual location of the point after rectification (Mather 2005: 87–89; Lillesand 2008 et al. 485–490).

During the georectification process the SPOT image was resampled to the new alignment and spatial resolution. When the spatial resolution stays the same, resampling is only done to make sure that the pixels are in straight lines, because they may be skewed in the rectification process. However in this case also the pixel size had to be changed to coterminous with the Landsat TM pixel. This was done with the nearest neighbor method where the pixel values of the georectified image are determined by the pixel values closest to them in the rectified but skewed image (Lillesand et al. 487–488).

5.7. Automated land cover classification

The classification process started with planning the classification system and recognizing spectrally separable land cover features from previous land cover classifications, aerial photographs of 2004 – 2005, GeoEye-1 images from 2009, Landsat TM image from 2009 (Figure 15). The originally planned classification system and its classes were modified multiple times during the process to see what actually was possible to classify from the two scenes, however constantly keeping the final use as deforestation modeling in mind.

The used classifications were created with a combination of unsupervised and supervised methods, which are the two most common automated spectral classification methods (Campbell 1996: 315; Lillesand et al. 2008; 545–547). In *supervised classification* the classifier directs the outcomes by creating example areas, called training sites for all predefined land cover features and computer classifies all the pixels in the image according to the spectral values of these training sites with mathematical algorithms, such as Maximum Likelihood, Minimum-Distance-to-Means or Parallelepiped. Maximum Likelihood classifying method quantifies the mean variance and covariance matrix of the created training areas and classifies other pixels according how well their properties match with these. Pixels probability of belonging to a class is calculated for each class in the classification and eventually the pixel is classified to the class with highest probability or as “unknown” if all the probability values are below defined threshold (Lillesand et. al 2008; 554–557). The training sites can be delineated in the field with GPS or visually from the used data with assistance of aerial photographs, high resolution satellite imaginary or maps (Campbell 1996: 329–334; Lillesand et. al 2008; 557–568). The classification outcome depends heavily on the quality of training sites, which should be as homogenous as possible to avoid internal spectral variations. Creating representative training sites is often time consuming process and requires knowledge about the geographical area, its annual changes and reference data from used sites (Lillesand et al. 2008; 557).

The advantages of the supervised method are that user can define the classes, make them to suite the study and ensure comparability to other classification schemes. The disadvantage on the other hand are that real land cover classes may not be spectrally unified and creating spectrally diverse training areas may bias the whole classification. Often there are also minor geo-rectification errors between primary and reference data, which might cause problems when delineating the training areas. Also if certain land cover features or classes are not represented in the training sites they are simply joined into spectrally most similar class, even though they would be far from this in the field (Campbell 1996: 327–329). These problems can be avoided by collecting training sites from all important land cover features, making sure that the sites are spectrally unified and combining features to more comprehensive land cover classes only after the technical classification process (Campbell 1996: 316).

In *unsupervised classification* the pixels are first clustered to predefined amount of natural spectral clusters based on their DN values and these clusters are attached to actual land cover classes. These clusters contain pixels which DN values are close as possible to each other and far as possible from other clusters (Campbell 1996: 317–319; Lillesand et al. 2008; 568–569). The clustering can be done with various algorithms, such as K-means and Iterative Self-Organizing Data Analysis Technique (ISODATA). In K-means approach user defines the amount of clusters he wants to identify. The algorithm then locates the defined number of cluster centers in the multidimensional data so that in the beginning the average distances between the clusters are equal. Each pixel is assigned to a cluster of which mean vector value is the closest to that pixel's spectral values. After this the mean vectors of clusters are re-calculated and the pixels are again assigned to the "closest" cluster. This iteration is continued until there are no more significant changes in mean vector values.

ISODATA follows the basic principles of K-means clustering, but in this algorithm the statistics of clusters are assessed after each iteration. If the distance between two clusters is lower than predefined minimum they are merged, while if the standard deviation is too high within single cluster it is split to two or the cluster is deleted if the amount of pixel within it is too low. The iteration stops when there are no longer significant changes in the cluster statistics or the maximum number of iterations is reached (Lillesand et al. 2008: 569–570). The unsupervised classes are solely spectral clusters and on itself they do not refer to any land cover classes, but they can be linked to land covers with reference data or field visits. Unsupervised classifications advantages are that it requires significantly less knowledge about the study area and possibilities of human errors are minimized. Although unsupervised classification appears more straightforward and cost-efficient than supervised classification, the process of connecting spectral clusters to real land covers requires large amounts of reference data and may turn out to be highly time consuming. It is especially problematic process when the classified image is so old that field assessments are not valid and reference data is scarcely available. Also the created clusters may not match the categories interested by the user or they are incomparable to other classification schemes (Campbell 1996: 317–319; Lillesand et al. 2008: 568–569).

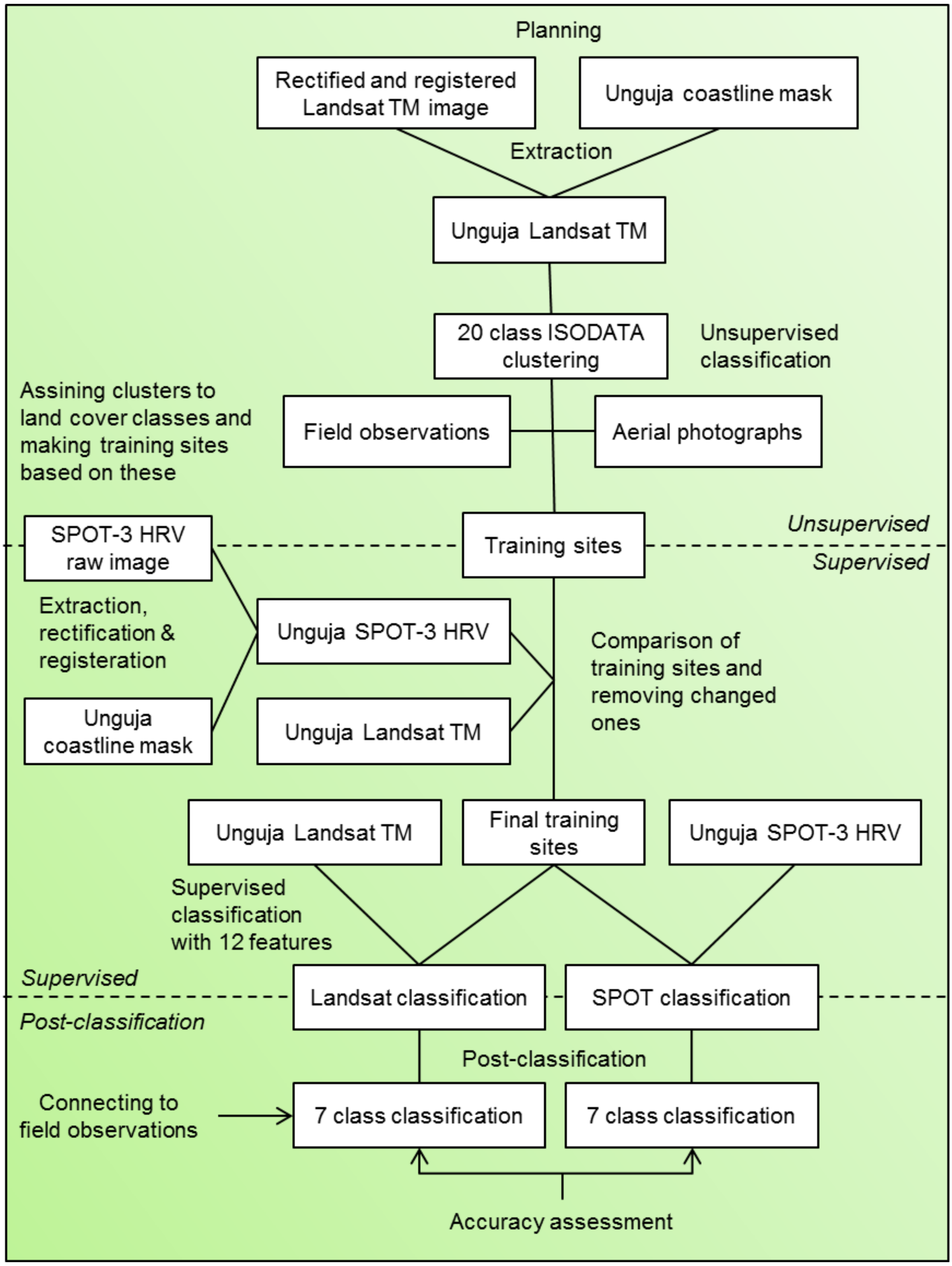


Figure 15. Flowchart of the main classification work phases.

After planning the already rectified and registered Landsat TM image was masked with the coastline of Unguja and classified with ISODATA unsupervised method to 20 clusters. These clusters were attached to land cover classes through field visits and visual assessment of aerial photographs. 18 out of 20 clusters were assessed in the field. The two clusters left out were urban and mangrove classes, so obviously recognizable from the aerial photographs that they were not considered to require any field assessments. The field observation points were chosen so that they were evenly divided between coral rag and deep soil regions, visited cluster patches were large enough to truly represent the land cover and that they were along the roads making the process efficient. The collected GPS points were imported to ArcGIS and overlaid on the created spectral clusters and the observations and photographs collected in the field were attached to the clusters. The ISODATA clusters, field visits and the aerial photographs helped to identify 12 spectrally separable land cover features.

To create coherent classifications between the 2009 Landsat TM and 1996 SPOT images the classification was turned from unsupervised to supervised and approximately 250 training sites were collected from these 12 primary land cover features. The training sites were based on the field visits, aerial photographs and knowledge from the previous research done in the island (Käyhkö et al. 2011; Fagerholm et al. 2012). In the beginning the training sites were collected only for the Landsat TM image, but later assessed and modified with the SPOT image. This was done by first matching the SPOT image histogram against the Landsat histogram to lower the contrast differences caused by different sensor calibrations (Wulder et al. 2008). Then the training sites were gone through one-by-one and they were modified or deleted if there were clear visual differences between the satellite images. After the primary training sites were collected the classification was tested and the spectral statistics of each classes training sites and the created classifications were visually assessed. The primary training sites were modified multiple times based on these assessments to make the training sites, their signatures and the whole classification as coherent as possible between the two images. In the end there were 203 training sites for the 12 spectral features, which were used to create signature files individually from both images. The final classification was done with maximum likelihood supervised classification method for radiometrically non-corrected images. The classified Landsat TM image included all the other bands except the thermal band and the SPOT image all the bands.

Because of large spectral differences between land cover features, the amount of classes was intentionally kept as high as 12 (Figure 16). For example there were four spectrally different forest features and combining them to one class would have made the range of DN values wide and risk of misclassification high (Campbell 1996: 327–329). Many of the different spectral land cover features were based on differences in near infrared values and these differences were estimated by calculating the mean

Normalized Difference Vegetation Index (NDVI) value of each class in 2009 classification before combining them together. NDVI value indicates the amount of live green vegetation within a pixel and it is calculated as:

$$NDVI = \frac{NIR - R}{NIR + R}$$

Where NIR is the near infrared band and R is the red band of the sensor (Lillesand et al. 2008: 464–466). After the NDVI values were considered coherent the classification was post-classified to 7 classes.

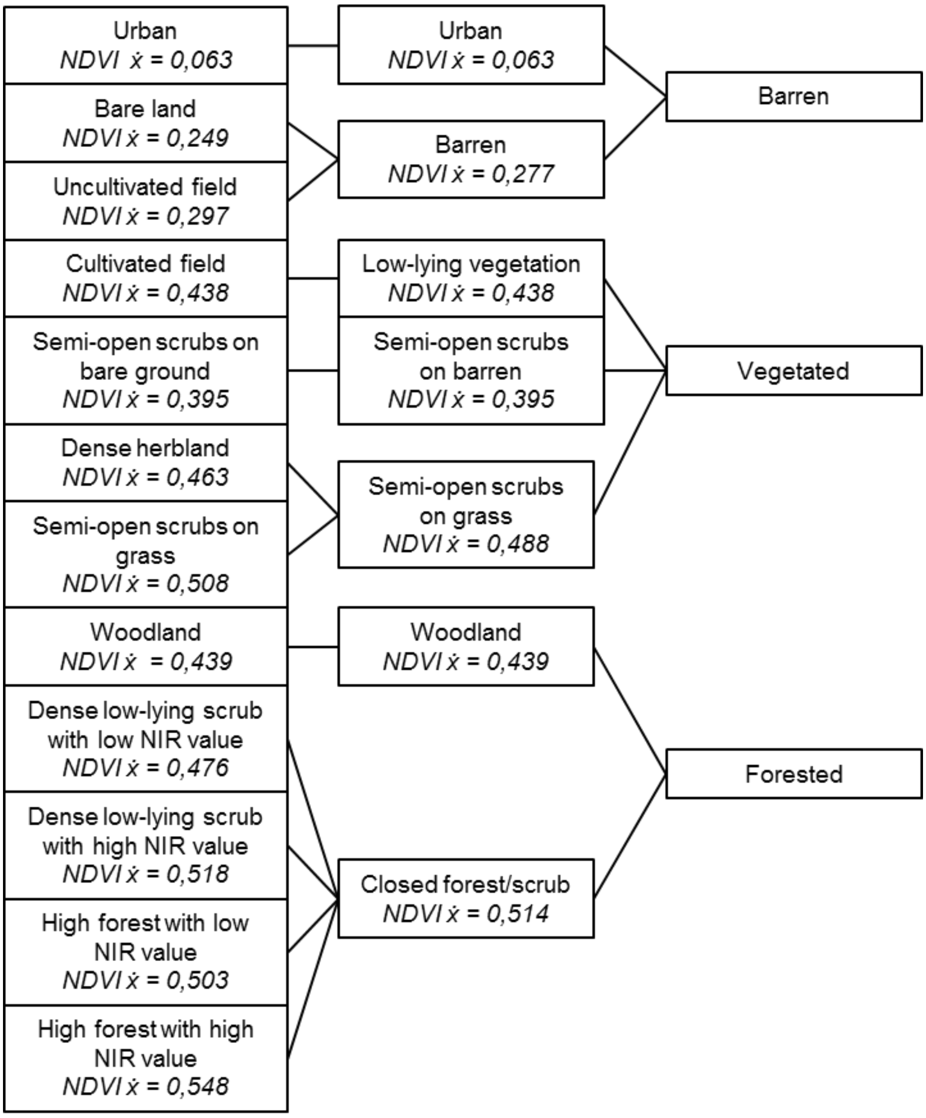


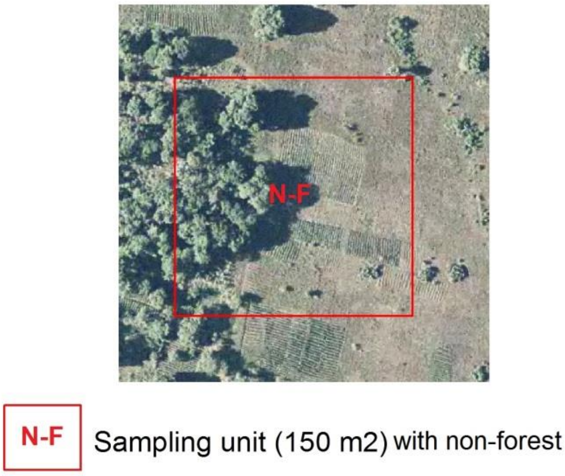
Figure 16. The original 12 spectral land cover features were post-classified hierarchically to seven more easily interpretable land cover classes and also handled as three different classes in the change detection. Mean NDVI values were also calculated for each class.

All classifications are eventually subjective and represent the interpreters' visions of the landscape. This makes replication or comparing classifications fairly difficult (Weiers et al. 2002). In change detection significant portion of changes can come from the minor differences between nominally same classes rather than from actual land cover changes (Weiers et al. 2002; Di Gregorio 2005: 7–10; Ahlqvist 2008). The subjectivity should be minimized in the classification process, but also qualitative descriptions (ea. land cover type and land use activities) and quantitative parameters (ea. percentage of canopy coverage and amount of tree layers) should be linked to the classes after creating them. These can be acquired from high resolution aerial images or with field observations and they have important value in describing the classes, underlining class differences, repetition of classification and in comparison with other classifications (Ahlqvist 2004). For these reasons, the created classes were described and parameterized with the collected field observation after the actual classification process. GPS from field observation sites were overlaid on the land cover classification of 2009 and connected to corresponding land cover class. This information was connected to the table created from observations. Key variables were chosen and most common, average and majority values were calculated. Some of the observations were left out from the comparison, because they were repetitions of previous points, did not contain all the needed information or had clearly changed after 2009. It should be remembered that the field observations were also used to associate unsupervised clusters to land covers and to delineate training sites, therefore the descriptions and parameters drawn from the observations and photographs taken are representing the classes better than randomly selected sites. In a way they are representing the ideality of the classes and not the objective reality. Giving objective descriptions or parameters to classes would have required an individual data collection campaign or post-classification interpretation of the aerial photographs, which was impossible when considering the time resources in hands.

5.8. Accuracy assessment

Aerial photographs and high resolution satellite images were used as reference data in this research and their systematic visual interpretation created the basis for accuracy assessment. In the accuracy assessment approaches the producer estimates visually the land cover in the sampling unit, based on created classification structure. This visually interpreted land cover is later compared against the class in automatic classification. The method relies heavily on the skills of the interpreter and visual interpretation elements are often set to decrease the subjectivity and to keep the outcome coherent (Lunetta & Lyon 2000; 7). Although some interpretation elements were helping the interpretation, their use was not systematic in this research.

Accuracy assessment was based on samples of the data pixels. There are various alternatives for sampling unit, but for satellite imaginary it is usual to overlay the created classification with a grid and estimate majority land cover within randomly chosen grid cells (Lillesand et al. 2008; 588). The size of the cells is determined by the original pixel size, but usually they are pixel blocks of 3 to 10 times the original size (Stehman & Czaplewski 1998, Lunetta & Lyon 2000; 7). Choosing samples from class boundaries is often avoided, because it is fairly tricky to draw an absolute line between two land cover types and inaccurate geo-rectification causes errors close to the edges. On the other hand majority of changes happen in these boundaries between classes, so their accuracy should be as precise as possible (Lunetta & Lyon 2000; 6–7, Lillesand et al. 2008; 587). In this study the assessment was done with a grid of 5 x 5 original pixel sized cells, which was overlaid on reference data and land cover majority was interpreted visually for each cell (Figure 17). The grid size was considered suiting the mosaic like structure of landscape where changes may happen even within short distances.



Sample ID	LCC in reference data	LCC in classification
1	Non-forest	Non-forest
2	Forest	Non-forest
…	…	…

Figure 17. Sampling cell overlaid on 2004/2005 aerial photograph representing non-forest in reference data and an example image of accuracy assessment attribute table.

High amount of pixels in Landsat TM image forces the user to sample the data rather than assessing all of the pixels. Even collecting the 0,5% sample, traditionally used in statistics, may be problematic if the study areas are large. Usually in the related literature minimum of 50 samples are chosen from each land cover class, but this amount is in relation to produced map (Congalton 1991). Some literature advices that that more samples should be chosen from the land cover classes that are in the focus, but this approach would distort calculation of some statistical figures like overall accuracy and KHAT statistics (Lillesand et al. 2008; 585–591). Rather the sample sizes should be in line with the proportional size of a class from the total land cover (Lunetta & Lyon 2000; 4–5, Lillesand et al. 2008; 588). So that all land cover classes would be well represented in accuracy assessment samples are often chosen with stratified random sampling, with limitations on the minimum and maximum amount of samples per class (Lunetta & Lyon 2000; 5, Lillesand et al. 2008; 588). Sampling cannot be completely random, because if randomly chosen samples cluster in each other vicinity the outcome is influenced by spatial autocorrelation. Congalton (1988) estimated that spatial autocorrelation influences neighboring land covers as far as 1,8 kilometers away from original sites, but this depends on the landscape at focus. To reduce the spatial autocorrelation sampling must be modified so that sampling sites do not cluster to the extent that closeness of other observations would influence the results (Lunetta & Lyon 2000; 6).

The land cover class in majority in the created classifications was calculated for each cell with “Zonal statistics” –tool in ArcGIS 10. Stratified random sampling of assessment units was performed by first adding the variance information of the classifications to the grid cells with “Zonal statistics” –tool. Only those cells that had variance value one, meaning that there were only one land cover class in these cells were chosen for assessment. This was done because it would be impossible to assess cells with high land cover diversity where majority could be reached with rather small amount of pixels (with 7 land cover classes the min. is 5/25 pixels). Random numbers were calculated in ArcGIS 10 “Field calculator” –tool for all remaining cells and they were set to ascending order. Classifications land cover majorities were removed from the attribute table and actual classification images were removed from the workspace to avoid the possibility of peeping. In situations where land cover change was obvious or clouds blocked the view, the land cover was not assessed. Altogether 361 sample units were checked from the Landsat and 358 for the SPOT classification. Sample size was approximately the recommended 0,5% of the total land cover (Lillesand et al. 2008; 588). The amount of sample units per class was proportioned to their size from total area. Assessed cells situated less than 450 m from other assessed cells were removed to reduce the errors caused by spatial autocorrelation. The threshold distance was set lower than suggested by Congalton (1998), since it was considered that the mosaic like structure of the landscape reduces spatial autocorrelation.

The observations collected from reference data were compared against the produced classification in error matrix. In this cross tabulation the reference data is represented in rows and classified data in columns, revealing which reference data observations fall in the same land cover class in produced classification and vice versa (Lunetta & Lyon 2000; 3, Lillesand et al. 2008; 585). Error matrix works as a baseline for multiple statistical measures used to assess the accuracy of classifications. Overall accuracy (*O*) is calculated as:

$$O = \frac{\sum_{i=1}^{r} x_{\mathrm{ii}}}{N}$$

Where all the correctly calculated observations ($\sum_{\mathrm{i}=1}^{\mathrm{r}} \mathrm{x}_{\mathrm{ii}}$) are divided with the total number of observations (N). Accuracy can also be calculated individually for all classes. Producer's accuracy refers to amount of correct observations divided by the total amount of reference observations in that particular class. This measures the accuracy of reference samples being correctly classified and how well certain land covers can be classified. User's accuracy indicates the probability that classified sample unit represents the same land cover in reality. It is calculated by dividing the correct observations with the total amount of sample units belonging to that particular class (Lunetta & Lyon 2000; 3–4, Lillesand et al. 2008; 585). Even completely random classification would provide an outcome with certain accuracy. KHAT statistics compares created classification to the average accuracy of randomly created classification. The KHAT (k) statistics calculated as:

$$k = \frac{N \sum_{i=1}^{r} x_{\mathrm{ii}} - \sum_{i=1}^{r} (x_{\mathrm{i+}} \cdot x_{\mathrm{+i}})}{\mathrm{N}^2 - \sum_{i=1}^{r} (x_{\mathrm{i+}} \cdot x_{\mathrm{+i}})}$$

Where *N* is the total number of observations in the matrix, *r* number of rows, x_{ii} is the sum of correctly classified observations in row *i* and in column *i*, x_{i+} is the total of observations in row *i* and $x_{+\mathrm{i}}$ is the total of observations in column *i.* K values vary from 0 to 1, where value 0 would refers to classification that is no better than randomly created one. K value can be interpreted so that if it is 0,76 the classification is 76% better than one created completely on random basis (Lillesand et al. 2008; 590 – 591).

The error matrix table was built manually in Microsoft Excel 2010 from the attribute table of sampling units with a SQL search of both classifications. The majority in the created classification was set to rows and in reference data to columns and the samples were cross tabulated. The cross tabulation was used to calculate overall accuracy and KHAT figure for the classifications and the user's and producer's accuracy for each classes.

5.9. Post-classification comparison

The change detection was done with post-classification comparison method, where images from different times are classified independently and then compared pixel-by-pixel or segment-by-segment (Coppin et al. 2004). The outcomes can be represented as change maps and transition matrixes. Change maps are visual spatial models of not only the changed areas, but also the persisted ones. Transition matrix is a statistical cross-tabulation of the classes in the two used classification. The transition matrix provides information about the quantity and direction of change class-by-class. Post-classification comparison is able to show how individual pixels and class areas changed through time. For example has the change been from forest to urban or what has been the area that changed from forest to agriculture (Coppin et al. 2004; Pontius et al. 2004; Lu et al. 2004; Lillesand et al. 2008: 595).

Transition matrixes can be used to calculate class and landscape level change statistics. *Net change* is the absolute value of area lost or gained by an individual class or the whole landscape. Although it is the key quantitative figure of change, it is essentially aspatial and thus not acknowledging spatial changes within the classes. For example the total amount of forests can keep stable as certain areas face deforestation which is counterbalanced by reforestation elsewhere (Pontius et al. 2004). In research of Mertens and Lambin (2000) from Cameroon over half of the changes were actually spatial changes within classes, while in local level case study from Zanzibar these changes covered over ¾ of total changes (Käyhkö et al. 2011). These spatial changes are called "swapping" and its measurement requires calculating such quantitative figures as *gain, loss, persistence, total change* and *swap* (Pontius et al. 2004). *Gain* refers to the amount of area that was not part of particular class in the first classification, but belonged to that in the second one, or more simply put, the area gained by the class in the time period. *Loss* is the change to opposite direction, area lost during the time period. *Persistence* measures the area of the class that did not change spatially. *Total change* sums gain and loss to calculate the total amount of area changed. *Swap,* measured by doubling the smallest gain or loss value or by subtracting net change from the total change, indicates the amount of location changes within the class. These figures can be also calculated for the whole landscape by summing class level values, however for swap and total change the summed values need to be divided by two, because loss in one class is compensated by gain in other classes (Pontius et al. 2004).

The change detection of SPOT 1996 and Landsat 2009 land cover classifications was performed with Erdas Imagine "Matrix union" –tool, which produces a new image where class value is determined by the combination of class values of the two used images pixel-by-pixel. Using seven classes created 49 types of change (7 * 7 = 49) (Appendix 2), which were cross-tabulated to a transition matrix. *Gain, loss, persistence, swap,* *total change* and *absolute value of net change* were calculated from the transition matrix based on methodology presented by Pontius et al. (2004). Some areas classified incoherently between the satellite sensors (ea. rubber plantations) were delineated with the DoSUP 2009 land use classification and reclassified to correct change class.

Previous research indicates that swapping and total changes are extremely high in the study area (Käyhkö et al. 2011). Also classification problems related to cross-sensor analysis and class semantics may increase the amount of swapping and total change (Pontius et al. 2004; Ahlqvist 2008; Wulder et al. 2008). Therefore it was seen important to generalize the original change detection outcomes. Pretesting indicated that pixels swapped frequently between rather similar land cover classes, such as *semi-open scrubs on barren* and *semi-open scrubs on grass*. Due to this the 49 types of change were post-classified to two generalized vegetation change classifications with 7 and 11 categories based on land cover classes and direction of change. The 7 classes included *reforestation, stable forest, deforestation, revegetation, stable vegetation, devegetation* and *stable non-vegetated,* while the 11 classes included also *forest improvement, forest degradation, vegetation improvement* and *vegetation degradation*. How the changes from 1996 land cover to 2009 land cover were classified to change classes is represented in Appendix 2. In the 7 class classification, which is now an called as "abrupt classification", *barren* and *urban* classes were considered as "barren", *low-lying vegetation, semi-open scrubs on barren* and *semi-open scrubs on grass* as "vegetated" and *woodlands* and *closed forests/scrubs* as "forested". Landscape change and transition matrix statistics were recalculated for these classes, causing *swap* and *total change* to lower to more acceptable levels.

Based on visual estimations and accuracy assessments it was noticed that changes from *semi-open scrubs on grass* and *woodlands* to *closed forests/scrubs* and vice versa were often misclassified. It was seen necessary to treat these classes separately as intervening categories not truly belonging to "vegetated" neither to "forested" in the three class division. In the 11 class classification, called as "gradual classification", *semi-open scrubs on grass* were handled as "degraded forest" class between "forested" and "vegetated", which meant that changes from "vegetated" to *semi-open scrubs on grass* were classified as *vegetation improvement* and as *vegetation degradation* to the opposite direction and changes from "forested" classes to *semi-open scrubs on grass* were considered as *forest degradation* instead of *deforestation* and as *improved forest* when it happened vice versa. Also the *woodland* class was approached in a similar manner, but so that changes from it to *closed forests/scrubs* was considered as *forest improvement* and *forest degradation* when happening to other direction and changes from it to *semi-open scrubs on grass* were considered as *forest degradation* and as *forest improvement* to the other direction. Using two different classification schemes reduces misclassifications. The classification with 11 is stricter when it comes to deforestation and reforestation and it represents forest change that has happened for absolute certainty, but in some cases it is too strict to detect all the happened changes. The 7 categories classification is able to detect these changes, but in some areas it is too loose and even subtle changes are marked as deforestation. The truth lies somewhere between these two and using them both allows leeway in estimations and instead of producing one deforestation figure this approach produced a range of deforestation.

Transition matrix and landscape change statistics were calculated for the abrupt classification, but the complexity of the gradual classification prevented these statistics to be calculated, therefore only the change maps and change class statistics were produced for this classification. When transition matrixes or landscape change statistics were used in the analyses it always refers to the ones created from the original 49 change class or the abrupt classification. Class level statistics were calculated for all of the 7 and 11 change classes. Focus was set on the forest changes, so the maps within the text were excluding other change classes, but the original products were included as appendixes and even though all classes were represented in class level statistics only forest changes were analyzed (Appendix 3). Deforestation, forest degradation, forest improvement, reforestation and net forest decrease figures were calculated for the whole 13 year study period. These figures were projected against the extent of stable, 1996, 2009 and all ever existed forests to estimate the magnitude of change and future developments of the forest cover. Changes were projected especially against stable and 2009 forests to estimate their stability, swapping and absolute deforestation in the future.

The absolute amount of forest decline is relational to the absolute amount of forests existing. In other words, when the forest stock declines so does the absolute amount of deforestation, though the rate of forest change stays unchanged (Puyravaud 2003). Therefore calculating the annual rate of forest change provides better estimations about the future of forest cover than simply calculating the absolute areas lost (FAO 1995; Puyravaud 2003). Based on FAO (1995) standards the annual rate of forest cover change (*q)* is calculated from the land cover data as:

$$q = \frac{A_2}{A_1}^{(1/(t_2 - t_1)} - 1$$

Where t_1 and t_2 are the times (year) of the forest estimates and A_1 and A_2 the estimates of these years. The figure can be turned to percentages by multiplying it with 100. The annual rate of forest change was used to project the current forest cover until 2050.

5.10. Kernel Density Estimation

The clusters of forest change can be mapped with various techniques, but Kernel Density Estimation (KDE) connected to Moran's I spatial autocorrelation calculation was used in this research to avoid subjective visual estimation of areas most changed. KDE was used to map the actual clusters, while Moran's I calculations determined the cut off distance for the KDE and measured spatial autocorrelation within this distance (Silverman 1986; ESRI 2012). Mathematically the kernel density ($f(x)$) for bivariate data is defined as:

$$f(x) = \frac{1}{nh^2} \sum_{i=1}^{n} K \left\{ \frac{(x - Xi)}{h} \right\}$$

Where n is the number of observations, h is the used cutoff distance, K is a kernel density defined as, (x) vector of x,y coordinates describing the location where the function is calculated subtracted with (X), the x,y vector location of each observations (i) divided with the cutoff distance (Silverman 1986; Seaman & Powell 1996). In more practical terms KDE calculates the occurrence of events within certain cut off distance (Figure 18). Different events or magnitude of events may be weighted differently (ea. reforestation 1, forest improvement 0,5, forest degradation -0,5 and deforestation -1). The value of events is not only relational to their weight, but also distance to them from the observing point influences, so that points further away get lower coefficient factor than those nearby. As a moving window method, the KDE starts from a point, calculates its kernel value in relation to other points, their weights and distances and the moves to the next point until all points in the data are gone through (Silverman 1986; Brinkmann et al. 2011; Grove 2011). Eventually it creates smoothed surfaces, where clusters of deforestation cover also pixel where deforestation has not taken place. The value for *point in focus* in example Figure 18 would be -2,5 if the occurrences and their weights would be summed together within the cutoff distance, however KDE also acknowledges the distance to the occurrences and therefore the value would be lower for the point, since the events with negative weights are closer than the ones with positive values. In the technical solution of ArcGIS 10 the final KDE value is eventually divided with the area created by the cutoff circle.

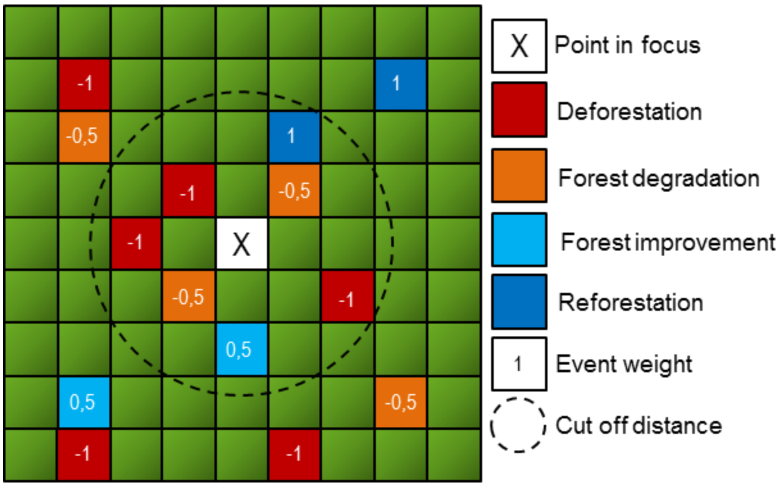


Figure 18.Visual representation of the Kernel Density Estimation. KDE is calculated by summing the events multiplied by their weight and distance from the point in focus in relation to the cut off distance.

The z-score from Moran's I calculations is considered to peak when spatial processes are most pronounced and this cut off distance should be used in distance based spatial analyses. The peak can be defined by testing different cutoff distance and comparing their z-score, however in some occasions there can be multiple peaks when different spatial processes function at different scales (ESRI 2012). In practice Moran's I calculation was done first by converting the abrupt classifications deforestation and reforestation pixels to points with Raster to Point –tool in ArcGIS 10. Moran's I calculations requires heavy mathematical data handling processes, which makes the use of whole study area impossible, therefore approximately 3 km^2 a test site was chosen from Muyuni area. The test site was chosen so that its spatial clustering would represent the clustering pattern of the whole study area as well as possible. Spatial autocorrelation of these points were calculated with various cut off distances (500 m, 750 m, 1000 m, 1250 m and 1500 m) with Spatial Autocorrelation (Moran's I) –tool and the distance with highest z-score were used in KDE.

Kernel densities are calculated from point patterns and therefore the forest change classes (reforestation, forest improvement, forest degradation and deforestation) in the abrupt and gradual classification rasters were turned to point data. The KDE were calculated for three different point sets with ArcGIS 10 "Kernel Density" –tool. Doing this provided information about different forms of deforestation and their patterns of spatial clustering. The first used point set were the deforestation and reforestation pixels in the gradual classification and this was considered to map areas of absolute changes from forested to other land covers or vice versa. The second KDE also included the deforestation and reforestation pixels from the gradual classification, but also the points of *forest degradation* and *forest improvement* with their own value weightings. This KDE mapped not only the key areas of absolute changes, but also the sites where degradation and improvement were present. The third data set included the *deforestation* and *reforestation* points from the abrupt classification and therefore mapped the key areas of change without taking any stands to the severity of changes or in other words making no distinctions between *deforestation* and *degradation* or *reforestation* and *forest improvement*.

Only the reforestation pixels with higher values than 0,000347226 and deforestation pixels with lower values than -0,000347226 were used in analyzes and mapping. This threshold was calculated by first extracting only the negative values from the third KDE raster and then classifying the remaining pixels based on quintiles to classes with 10% share from the landscape. The dividing value between the most deforested 10% of pixels and the rest was 0,000347226, which was set as the threshold for deforestation hotspots. The value was turned negative to determine the boundary for reforestation hotspots so that these would be comparable to the deforestation hotspots. However the reforested pixels had originally smaller values and therefore the hotspots of reforestation no longer represented 10% of the landscape. Eventually the deforestation and the reforestation hotspots from all the three point sets were mapped over the generalization of the study area map.

5.11. Sub-area, distance and multivariate regression analysis

In this research forest changes were linked to environmental factors with three quantitative approaches. Firstly the island was divided into four different forest sub-types based on their soils and government status and the forest dynamics differences between these types were studied. Secondly the influence of the vicinity of Zanzibar Town, main roads and coastline were studied by relating Euclidean distant measurements to values of forest change at the level of the entire island. Thirdly the forest types, distance measurements and other environmental factors were used as explanatory variables in regression analyses which modeled deforestation.

Driving forces having essential influence in one landscape may have only minor impact on another. Therefore it is important to understand the inner variations of the studied area and on demand it should be divided to smaller homogenous units. These divisions should be well justified based on literature, previous research or empirical testing (Mertens & Lambin 1997; Serneels & Lambin 2001). However dividing to too many small units is time consuming and unable to achieve generalizations (Serneels & Lambin 2001). In the case of Unguja it was seen that the deforestation processes and causes are highly different for the western agroforest and eastern indigenous forest regions. As a starting hypothesis it was assumed that the agroforest as important component of the agricultural system are rarely deforested without serious cause, while on the other hand the indigenous forests as sites of swiddening, fuel wood and building material collection are more prone to deforestation (Käyhkö et al. 2009; Fagerholm 2012: 31–37). Spatially the division between these forest types was done based on the soils in the physiographic map of Hettige (1990) digitized by Klein (2008), in a manner that the forests on deep soils were considered *agroforests* and the ones in coral rag as *indigenous forests/scrubs*. However there are significant differences even within these areas mainly caused by different government forest management actions. Therefore the *government protected forests* and *government forestry forests* were separated from the two main categories with official forest station delineations of DFNR.

Transition matrixes, landscape change statistics and change class statistics were calculated for each forest type. Some figures like the absolute areas of forests in 1996, 2009 and stable forests, each forest types proportional shares from these forests covers, the share of these forest covers from whole forest type area and share of stable forests from 2009 forests within the forest type were calculated. The average deforestation rates for each forest type were calculated with the same FAO equation than the rates of entire Unguja and the differences of gradual class distributions were examined by putting them on column charts. Landscape change statistics for each forest type were also calculated. However it is impossible to compare total change, swapping, stable forest and net deforestation figures of different forest types directly since the baseline situations of forest cover have been completely different, therefore the figures need to be proportioned against the share of forest cover within the forest type and reflected against the proportion of the same processes at the level of entire Unguja. This is done with the following equation:

$$X = \frac{a/b}{c/d}$$

Where the proportional share of any change process (total change, swap, net deforestation, stability) within class forest (X) of any forest region (agroforest, indigenous forest, government protected forest, government forestry forest) is calculated by first dividing the change process within the class forest of that forest type (a) with the total amount of class forests existing during the time period within the same forest type (b). This outcome is then divided by the outcome of the same change process in Unguja's forests (c) divided with the total amount of forests existing during the time period in Unguja (d). The outcomes can be interpret so that when the value is 1 the change process in that forest type has happened with the same rate than in entire Unguja when the baseline situations of forests have been taken into account.

In the second approach the relationships of forest changes and distance measurements were studied nonlinearly. The used method used was roughly similar to the one used by Mertens & Lambin (1997). Main and secondary roads, coastline and border of Zanzibar Town in 1996 were chosen as the sources of the Euclidean distance measurements. Buffers of 60 meters (2 pixels) were created around all these elements to exclude errors caused by georectification. Outwards facing distance rasters were calculated from the buffers and then reclassified after every 150 meters until 3 kilometers. Percentage shares of deforestation, reforestation, stable forests and net deforestation from the total forest cover of 1996 were calculated for each 150 meter zone from the gradual classification. These shares were proportioned against the average of the entire Unguja as was done in the forest type analyses, with the exception that the baseline situation was not all the forest cover ever existing, but the cover in 1996. The proportional shares were visualized as line diagrams, which follow the same formal logic as the change ratios of different forest types: when the value is over 1 the change process has been more common than averagely on the island, while if the values are lower than 1 the process has been less common and the negative values of net deforestation refer to forest gain instead of loss.

Linear trendlines and their coefficient of determination (R^2) were calculated from the net deforestation ratios of each distance measurement. These line diagrams and their trendlines enable better understanding of the linearity, nonlinearity and range of influences (cutoff distance) of each distance measurement used. This knowledge helps to interpret the regression modeling outcomes, which rely solely on linear relationships between variables (Verburg et al. 2004a; Metsämuuronen 2008: 119–121). However, also the cumulative percentages of forest cover were calculated for each distance measurement to allow understanding of their magnitude of influence at the scale of entire Unguja.

In the third approach linear regression analyses were used to determine the power and direction of individual variables on deforestation and to estimate the deforestation probability of still existing forests. Binary logistic regression was chosen as the used method from the family of regression analyses, mainly because of two reasons. Firstly it allowed using binary dependent variable of deforestation/stable forest instead of continuous variable of net deforestation. The idea was not to find out why certain cells deforest more than other, but to examine what separates cells of heavily declining forest cover from those of extremely stable cover. Secondly its mathematical basis was easily understandable, which makes the interpretation of outcomes and creating future scenarios easier (Metsämuuronen 2008: 114–126). Binary logistic regression follows the following equation:

$$\ln\left[\frac{P(Y=1)}{1-P(Y=1)}\right] = a + b_1x_1 + \cdots + b_\mathrm{n}x_\mathrm{n}$$

Where logarithm ($\ln$) of the probability of the dependent variable being 1 [$P(Y=1)$] divided with the same probability subtracted from 1 is equal to explanatory variable (b) multiplied with the regression coefficient (x) added to regression constant (a) (Metsämuuronen 2008: 116).

The regression modeling was done with aggregated cells. This generalization was seen to reduce the errors caused by misclassification of individual pixels and it also decreased the data handling and calculating capacity required. Using spatially too detailed dependent variable could have prevented from understanding the grand lines of the deforestation process. The original deforestation data was aggregated to cells of 300 m^2 (ten times the original Landsat TM pixel) and turned to vector format. The created vector database included altogether 17 958 cells, which were in spatial alignment with the original Landsat TM pixels. The information from land cover classifications of 1996 and 2009 and gradual change classification were fed into database class-by-class, so that each cell has percentage information about the forest cover in 1996, 2009, deforestation, reforestation and their combined net deforestation, which were then used to calculate the binary dependent variable (Table 9).

Table 9. Example of the regression database

	Cell ID	1	2	3	4	5
Origin of dependent variable	1996 forests (%)	10	95	74	20	10
	Deforestation (%)	10	31	27	20	10
	Reforestation (%)	0	1	7	0	25
	Net deforestation (%)	0	68,4	73	100	-50
Dependent	Deforestation (1/0)	0	1	1	0	0
Explanatory variable	Distance to coast (km)	2,0	0,5	3,1	1,4	5,0
	Distance to Zanzibar Town (km)	1,5	3,4	2,1	1	15
	Distance to roads (km)	4	2,5	0,2	1,4	5
	Building kernel density	245	95	145	523	16
	GPS (1 = yes / 0 =no)*	1	0	0	0	0
	GFA (1 = yes /0 = no)**	0	0	1	0	0
	Mean NDVI	0,354	0,489	0,648	0,214	0,712
	Mean elevation	10,2	24,5	58,6	1,3	7,5
	Soil (1= deep soils, 2= coralline)	1	2	1	2	2

*) Government protection status **) Government forestry status

Deforestation in gradual classification was considered spatially more accurately modeled and therefore it was chosen as the basis for the dependent variable calculation. Dependent variables were created based on following equation:

1) Deforestation term:

$$x = \frac{a-b}{c} \geq 50 \ \cap \ \frac{c}{d} \geq 50$$

0) Stable forest term:

$$y = \frac{a-b}{c} \leq 10 \ \cap \ \frac{c}{d} \geq 50$$

The dependent variable gets value 1 when the term of heavily declining forest cover (x) is achieved. This state was defined as a situation where deforestation (a) subtracted with reforestation (b) and divided with forest cover of 1996 (c) was over 50 (%) and (∩) the forest cover of 1996 was over 50 (%) from the total cover of the cell (d). In other words, the forest cover must have been over 50% of the cell in 1996 and at least 50% of this must have been deforested during the study period. The term for stable forests (y) was defined as a situation where cell had net deforestation lower than 10% during the study period and forest cover over 50 (%) in 1996. Altogether there were 662 cells (3,69%) that fulfilled the deforestation rule. The effects of spatial autocorrelation were reduced by manually removing all those cells that were closer than 1,5 km (3 cells) from other cells of the same class, which lowered the amount of deforestation cells to

186 (Congalton 1988). Originally there were 4583 cells (25,52%) that fulfilled the stable forest rule, but they were reduced to 187 to avoid overrepresentation in the modeling (Appendix 4). This was done by assigning random numbers to all the *stable forest* cells and choosing the 187 cells with the lowest random number, which were at least 3 cells away from each others.

The selection of used explanatory variables was based on related literature. Veldkamp & Lambin (2001) underline that the scale of study influences variable selection so that social and accessibility factors should be used at household level, topography and agroclimatic conditions at landscape level and climate, political-economy and demographic variables at regional or national level. The right variables for Unguja were seeked from other landscape and regional level studies from Africa and Tanzania. These studies have used such variables as elevation, slope, soil, land cover, vegetation, protection status, climate suitability, distance from coastline and accessibility to major cities, villages, markets, roads, rives and agricultural areas (Prins & Clarke 2007; Milledge et al. 2007: 6; Ahrends et al. 2010; Tabor et al. 2010; Swetnam et al. 2011). Some of these variables were left out from the modeling for various reasons. Even though there are differences in the rainfall between eastern and western parts of the island, climate was left out from the analyses because patterns caused by it are already present in the distribution of forested cells and correlates with soil data. The influence of river vicinity was not included, because of the small amount of rivers. Also the slope was not used as explanatory variable because lack of data. In the end altogether 9 explanatory variables were chosen for the regression analysis: distance to coast, distance to Zanzibar Town, distance to main and secondary roads, kernel density of buildings, mean NDVI, mean elevation, soil, government protected areas and government forestry areas (Table 10). Majority of variables are in continuous form, except for soil (1= deep soils, 2 = coralline soils), government protected areas (0 = no, 1 = yes) and government forestry areas (0 = no. 1 = yes). The variables were collected from various sources and modified to fit the regression analysis.

Table 10. The regression analyses variables, their sources and types.

Variable	Source	Variable type
Distance to coast	DoSUP 2009 topographic database	Continous (km)
Distance to Zanzibar Town	SPOT 1996 satellite image	Continous (km)
Distance to roads	DoSUP 2009 topographic database	Continous (km)
Building kernel density	DoSUP 2009 topographic database	Continous (kernel density)
GPA (1/0)	Government forest boundaries (DFNR)	Dichtomous (1/0)
GFA (1/0)	Government forest boundaries (DFNR)	Dichtomous (1/0)
Mean NDVI	SPOT 1996 satellite image	Continuous (0 - 2)
Mean elevation	DEM derived from contours of DoSUP 2009 topographic database	Continous (m)
Soil	Phyiographic map of Hettige (1990)	Dichtomous (1/2)

Multicollinearity of explanatory variables may flaw the modeling, therefore it should be tested before the actual modeling and the correlating variables should be dropped from the final model or combined as one (Serneels et. al 2007). Some regression methods include automated testing of multicollinearity, however this is not included into the binary logistic regression and therefore it was tested with Pearsons Bivariate correlation analysis with two-tailed test of statistical significance. Pearsons correlation analysis is the most common method to determine the correlation between at least two variables. The variables should be at least in interval scale and their dependence should be linear (Metsämuuronen 2008: 11–23). The Pearsons correlation coefficient (r) is calculated as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n s_x s_y}$$

Where n is the amount of number pairs of x_i and y_i, s_x and s_y are the standard deviations of the variables x and y and $\bar{x}$ and $\bar{y}$ are the averages of variables x and y. The outcomes can be interpret so that if r is 0 there is no correlation, if it is -1 there is negative correlation and value 1 refers to positive correlation. The outcomes of the correlation analyses were taken into account so that if variables had correlations over 0,7 or under -0,7 one of the variables was removed and correlation over 0,4 or under -0,4 were marked and considered when modifying the regression models or analyzing the outcomes.

Normal binary logistic regressions were calculated individually for each explanatory variable and conditional stepwise binary logistic regressions were used when combining all the explanatory variables to a single model. These multivariate models were calculated for the entire island and separately for coral rag and deep soil regions.

The island was divided between landscape regions to improve the modeling and to study the differences between these regions. The division was rough and some enclaves of deep soils were left within otherwise coral rag landscape and vice versa. The division was done in this rough manner rather than solely based on soils, because it was seen that the remaining enclaves are more influenced by the surrounding landscapes rather than areas of similar soil conditions kilometers away. The multicollinearity analyses were also done independently for the explanatory variables in each landscape region.

The outcomes of regression analyses were analyzed based on their regression coefficients, amount of correctly classified dependent variables, Wald statistics and Nagelkerke R^2. The regression coefficients are produced while calculating the regression equation introduced previously. The regression equation is tested against the used dependent variables to see how many of these variables are correctly classified in both classes and altogether based on the regression equation. Wald statistics is used to estimate the significance of individual explanatory variables in multivariate regression models. It is calculated as:

$$Wald = \left(\frac{\beta}{s.e.}\right)^2$$

Where the regression coefficient β divided with standard error $s.e.$ is squared (Metsämuuronen 2008: 118–119). Nagelkerke R^2 is a pseudo-R estimate implying the overall usefulness of the model. It is a modification from Cox and Snell R^2, which is calculated as:

$$R^2 = 1 - e^{\frac{2(LL(B)-LL(0))}{n}}$$

Where the Neper unit e (≈2,718) is squared with the log-likelihood of the created model $(LL(B)$, subtracted with the log-likelihood of antimodel $(LL(0)$, multiplied by 2 , divided with the sample size n and then subtracted from 1. In theory this calculation should reach 1 when the model explains all the dependent variables correctly. However in practice it never reaches 1, but Nagelkerke has modified the calculation so that it does reach zero when the model is perfect. Nagelkerke R^2 is calculated as:

$$\tilde{R}^2 = \frac{R^2}{1 - e^{\frac{2LL(o)}{n}}}$$

Where Cox and Snell R^2 is divided with the term $1 - e^{\frac{2LL(o)}{n}}$ dictating the maximum value possible to achieve with the model. There for Nagelkerke R^2 is able to directly tell how much the model is able to explain from the variations in the observations (Metsämuuronen 2008: 123).

Purpose of the regression analyses was not only to determine the direction and power of each independent variable, but to use the outcomes also for creating future scenarios. These projections were done in static “business as usual” manner, meaning that changes in rates, causes or spatial distribution of deforestation were not taken into account (Verburg et al. 2006: 119; Ahrends et al. 2011). Except that the new planned conservation area of Muyuni-Jambiani in the South district was included into the scenario building. The modeling was fuzzy in a sense that only the risk of deforestation ranging from 0 to 100% was visualized and the precise location and time of deforestation were left out from the scenarios (Mertens & Lambin 1997; Veldkamp & Lambin 2001). The modeling was done by using the explanatory variables and regression coefficients explaining the change between 1996 and 2009 to create regression equations for the forested cells of 2009. This was done independently for the coral rag and deep soil regions. Eventually it was only 3 variables that explained deforestation in the coral rag region and 4 variables in the deep soil. Some of the original variables just did not explain the variations or the outcomes were so bizarre that it was expected that the variables modeled something that were not really connected to them. For example “mean elevation” had really limited explanatory value in the individual models and in the multivariate analysis from the coral rag, but in deep soil region it suddenly became an important variable. The elevation differences in the island are really minor and therefore it was not considered plausible that increasing mean elevation would cause deforestation. Also it was seen distorting to include the “government forestry forests” into the modeling as the deforestation in these areas are caused by management actions that hardly fit to mathematical models. These variables were removed and the regression coefficients of the remaining variables were recalculated for the predictive modeling.

6. Results:

6.1. The distribution of forests and other land covers in 2009

6.1.1. General landscape

In 2009 approximately 38% of the island is covered by different kinds of forest vegetation, such as moist evergreen tropical forest, dry coral rag forest, forest plantations, agroforests, dense thickets and scrublands and semi-closed woodlands (Table 11 & Figure 19). Half of the area is covered by other forms of vegetation, such as semi-open scrublands and low-lying grasslands and the remaining 12% are non-vegetated, including urban areas, cleared agricultural fields, salt marshes and beaches. The overall accuracy of the classification is 82,5% and Kappa 0,784 (Table 12). Though, some classes such as low-lying vegetation, semi-open and woodland are more unreliable than the others.

Table 11. Land cover statistics of Unguja in 2009 based on classification of Landsat TM (01.07.2009) image.

2009 Land cover classification			
Class	Pixels	Area (km$^{2)}$	Percentage of landscape
Barren	103833	93,5	6,2
Urban	101033	90,9	6,1
Low-lying vegetation	186522	167,9	11,2
Semi-open scrubs on barren	286544	257,9	17,2
Semi-open scrubs on grass	352667	317,4	21,2
Woodland	138078	124,3	8,3
Closed forest and scrub	497489	447,7	29,9
Total	1666167	1499,6	100

Table 12. Error matrix with accuracy assessment figures for 2009 land cover classification of Landsat TM image from 01.07.2009. Aerial photographs of DoSUP from 2004 – 2005 and GeoEye-1 high resolution satellite images from 2009 were used as reference data.

		Reference data								
	Class	**B**	**U**	**LV**	**SB**	**SG**	**W**	**CF**	**Total**	**User's accuracy**
	B	32							32	100,0
	U		20						20	100,0
	LV			23	2	10			35	65,7
Classified data	**S-O B**	3		3	37	8	6	2	59	62,7
	S-O G			1	7	64	2	3	77	83,1
	W						22	5	27	81,5
	CF					5	6	100	111	90,1
	Total	35	20	27	46	87	36	110	361	
	Producer's accuracy	91,4	100,0	85,2	80,4	73,6	61,1	90,9		
	Overall accuracy	82,5								
	Kappa	0,784								

B = Barren, U = Urban, LV = Low-lying vegetation, SB = Semi-open scrubs on barren, SG = Semi-open scrubs on grass, W = Woodland, CF = Closed forests and scrubs

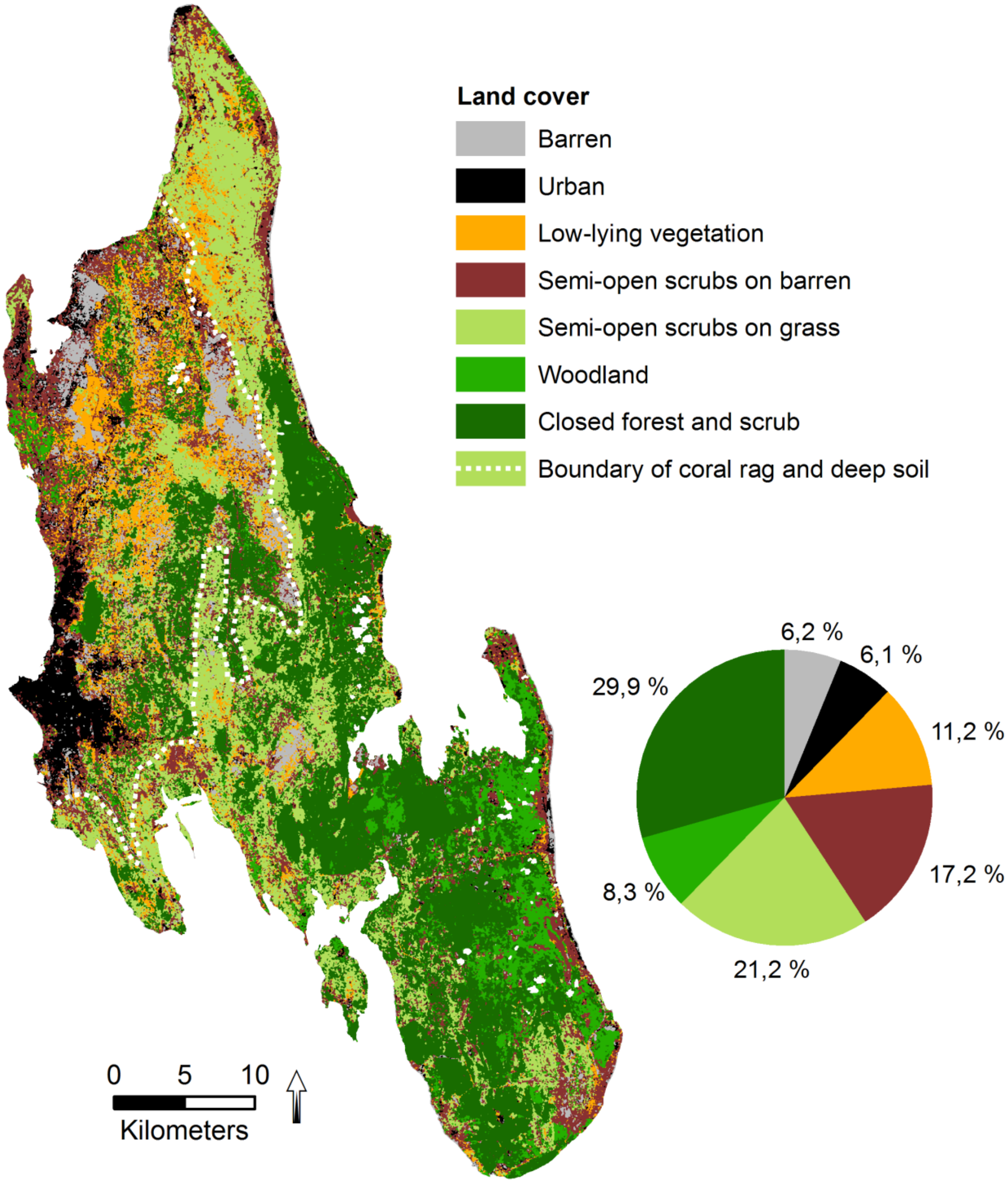


Figure 19. 2009 land cover classification with 7 classes classified from Landat TM images (01.07.2009).

There is a clear distinction between the eastern and the western sides of the island. The eastern coral rag region consists of large unified patches of closed forests and scrubs, woodlands and semi-open scrublands, while the western deep soil region is more a mosaic of small patches from various land cover classes spreading around villages or agricultural areas. Although there are some larger unified patches of urban, barren, low-lying scrubs and semi-open scrubs in the west. In general the deep soil region is structurally more a minute mosaic landscape of open agriculture, agroforestry and settlements, while coral rag is dominated by various forms of indigenous vegetation broken only occasionally by shifting cultivation or settlements.

6.1.2. Non-vegetated land

Relatively small portion (12,3%) of the total terrestrial landscape is non-vegetated, which included classes *barren* (6,2%) and *urban* (6,1%). Based on interpretation of aerial photographs uncultivated fields, beaches, sand taking areas, salt marshes and coral outcrops are classified as *barren* (Figure 20). Based on the producers' accuracy some semi-open scrubs on barren grounds are also classified to this class. Spatially the barren lands is scattered all around the island, but biggest concentrations are the salt marshes in the northwest and the uncultivated open agricultural areas in the middle parts of the island. *Urban* areas have 100% accuracy in all figures. Zanzibar Town and its conjunctions cover approximately 40% of the class, and also some of the smaller towns (Nungwi, Paje, Mkokotoni and Chawka) create relatively large unified urban patches.

6.1.3. Vegetated land

Vegetated land cover approximately half (49,6%) of the landscape and is divided between *low-lying vegetation* (11,2%), *semi-open scrubs on barren* (17,2%) and *semi-open scrubs on grass* (21,2%). *Low lying-vegetation* class consists mainly of cultivated and fallow fields, grasslands and wetlands. These land covers are clearly distributed more on the western side of the island, on the fertile deep soils and even the few field observations gathered from these sites showed higher moisture and looseness of soils (Table 13). Often these areas are under agricultural and livestock land uses, though there are also some natural wetlands and forest fire sites. In deep soil region low-lying vegetation patches are part of the larger rural landscape mosaic, while in coral rag these appear mainly in small swidden fields. The producers' accuracy (85,2%) is good, but the user's accuracy (65,7%) is rather poor and in both cases mismatches are caused mainly by other vegetated classes.

Table 13. Chosen field observation variables collected connected spatially to class of the site.

	Low-lying vegetation	Semi-open scrubs on barren	Semi-open scrubs on grass	Woodland	Closed forest/scurb
Observations sites	5	19	18	5	32
Land cover	Open cultivation (2), grassland (2), semi-open scrubland on barren (1), coconut plantation (1)	Semi-open coral scrub (9), coconut plantation (4), forest (2), grassland (2)	Semi-open forest/scrub (7), coconut plantation on grassland (6), low-lying forest (6), fallow (2)	Medium high forest (3), semi-open scrubs (2)	High forest (13), medium high forest (10), low-lying scrubs (2), fragmented forest (3), coconut plantation (1)
Canopy coverage	Open (3), semi-open (2)	Semi-open (11), closed (3), open (2)	Semi-open (11) to closed (8)	Closed (2), semi-open (1)	Closed (23), semi-open (3)
Tree/scrub layers [1]	0,6	2,1	2,1	2,3	2,6
Sites with low, medium and high trees (%) [2]	N/A	47 - 6 - 47	32 - 26 - 42	N/A	13 - 33 - 54
Sites with low, medium and high trees (%) without coconut [2]	N/A	80 - 0 - 20	43 - 36 - 21	N/A	13 - 35 - 52
Soil type	Loose soils (4), coral rag (1)	Coral rag (6), semi-coral (5), sand (4), loose soils (3)	Loose soils (9), semi-coral (8), coral rag (2)	Semi-coral (4)	Semi-coral (12), coral rag (8), loose soils (8), sand (1)
Bare groundlayer (%) [1]	N/A	38,5	11,5	N/A	68
Moisture conditions (dry = 1 & wet = 10) [1]	5,2	2,6	3,8	2	3,5
Land use	Permanent and shifting agriculture, grazing, coconut plantation	Cultivation, agroforestry, coconut plantation, spice farming, sacred site, garbage dumb, construction, conservation	Grazing, firewood collection, tree planting, coconut plantation, fallow, shifting cultivation, coral stone extraction	Forest conservation, casuarina plantation	Forest conservation, teak, pine, casuarina and coconut plantations, coral stone extraction, firewood collection, coppicing
Past land use	N/A	Shifting cultivation, gravel taking	Forest fires, shifting cultivation, acacia plantation	N/A	Shifting cultivation

[1)] Average [2)] Low = < 3 m, medium = 3 – 5 m, high = > 5 m

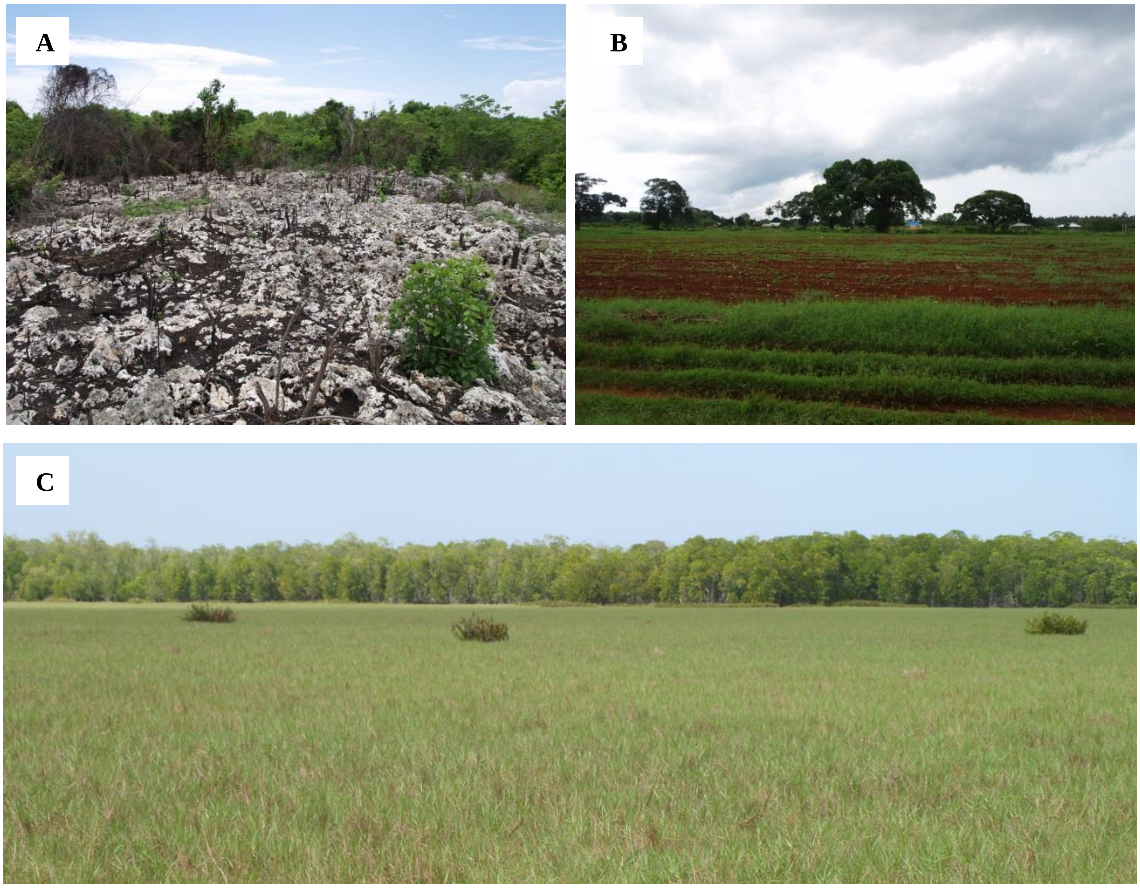


Figure 20. A) Coral outcrop revealed by burn-clearing. B) Open agricultural field in Bambi ward cultivated by military personnel. C) Wetland in Jozani National Park. Photos: Niina Käyhkö, 2011.

Semi-open scrubs on barren are distributed quite steadily between coral rag and deep soil regions. The biggest spatial concentrations are the sparse agroforest areas in the deep soil region, the couple years old swidden fields in the coral rag region and the naturally semi-open scrub areas on bare coral rag. The class is part of a swiddening continuum ranging from barren back to forest in coral rag and part of rural landscape mosaic in the deep soils. 19 field observation sites belonging to this class are mainly semi-open coral scrubs on grounds with limited or now undergrowth, but also some sparse coconut plantations, forests and grasslands appear (Figure 21). Average amount of bare ground layer is 38,5%, which is significantly higher than for *semi-open vegetation on grass*. This intensive reflectance from bare ground eventually separates these two semi-open classes. The canopy cover is mainly semi-open and 80% of the sites have highest tree layer being less than 3 meters high when coconut trees are left out from the calculations. Land use in the field observation sites is highly connected to agriculture, but also some urban activities like dumbing of garbage and construction are present. The users' accuracy (62,7%) is low and it seemed like all possible land cover types, except urban are misclassified as *semi-open scrubs on barren*. Producers' accuracy on the other hand is rather good (80,4%) and misclassifications are caused only by other vegetated classes. There are misclassifications to both directions between the semi-open classes.

Figure 21. A) Semi-open scrubland on barren ground in Matemwe shehia in northeast of Unguja. B) Semi-open scrubland on grass in Unguja Ukuu shehia in central Unguja. C) Mixture of scrubs and herbs close to village of Nungwi in north of Unguja. Photos: Niina Käyhkö, 2011.

Semi-open scrubs on grass cover almost the entire coral rag region, which is not already covered by forested classes. Based on the field observations and visual estimation of the aerial photographs and high resolution satellite images, the patches in this class are mixtures of grasses, herbs, scrubs and trees in various forms. From the 18 field observations majority are either semi-open scrubs, coconut plantations on grasslands or low-lying scrubs, separated by the grass or herb covered ground layer from the other semi-open class. The soils are looser than in the other semi-open class, allowing binding of moisture and vegetation growth in the ground layer. Majority of field observation sites have semi-open canopy layer (58%), but there is also quite a lot of sites with closed canopy (42%), which could refer to significant classification error. Land use activities are agriculture related, but also some forest land uses, such as firewood collection and tree planting are present. The class creates large unified patches on the coral rag, but on the deep soil these are smaller and often fragmented by other classes. Largest unified patches are found from the northern tip of Unguja and from the western parts of coral rag region. Although, the high amount of field observations had different land cover, the class has relatively good users' accuracy (83,1%) and mediocre producers' accuracy (73,6%). In the accuracy assessment the misclassifications came mainly from forested and other vegetated classes, and only 6% of the samples assessed as *semi-open scrubs on grass* from the reference data were classified as forests, even though the amount of sites with closed canopy layer was as high as 42% in the field observations.

6.1.4. Forested land

Majority of the *woodlands* are situated in the southern coral rag region, either as natural woodlands, ea. two circular areas within Jozani National Park (Figure 22) or as forested swidden sites, which are too high and vivid for semi-open classes, but not closed or high enough as *closed forests/scrubs*. Also some slightly open agroforest areas in the deep soil region are classified to this class. Field observations from woodland sites are limited and able to tell only minor details of this class. The land covers in the field observation sites are divided between medium high closed forests and semi-open scrubs and the amount of tree/scrub layers is slightly higher than for the semi-open areas. The soil type is semi-coral in all of the observation sites. The land use activities are linked to forest conservation or forestry and there is not any present evidence of agricultural activities. The class has a mediocre (81,5%) users' accuracy, but the producers' accuracy is low (61,1%). There are misclassifications to both directions between the *woodlands* and *closed forests/scrubs* and some *woodlands* are misclassified as semi-open scrubs.

Figure 22. Natural woodland in middle of Jozani National Park. Photo: Niina Käyhkö, 2011.

Closed forests and scrubs are largest of the individual classes and included moist evergreen tropical forest, dense and high thickets and scrublands, dry forests on coral rag and dense agroforests (Figure 23). Differently from *semi-open* and *woodland* classes the canopy coverage is always closed. The amount of average tree layers is highest from all of the classes and the class has largest percentage of trees over 5 meters, however in 48% of field observation sites the highest tree layer was less than 5 meters in height. The soils vary from coral rag to loose soils and land use activities are related to forest uses, extraction of soil and conservation. Opposite to open areas forests and scrubs concentrate on the eastern coral soil region of the island, where soils do not allow permanent agriculture. The forest in the coral rag region can be seen as one continuous area reaching from Kiwengwa to Muyuni or as three separate patches. These patches are surrounding forest conservation areas of Kiwengwa-Pongwe, Jozani-Chawka Bay and Muyuni. Two of these areas are joined to one unified patch covering almost 10% of the whole island. Between the main forest areas there are open or semi-open areas causing fragmentation of habitats. In the deep soil region forests are mainly dense agroforests, except for the Masingini forest conservation area just north from Zanzibar town. The class had high users' (90,1%) and producers' (90,9%) accuracies, though some semi-open scrubs and woodlands are misclassified as closed forests and some forests are misclassified to these classes.

Figure 23. A) Wet evergreen forest close to Masingini forest. B) Dry coral rag forest close to Jozani National Park. Photos:. Niina Käyhkö, 2011.

6.2. Forest change between 1996 and 2009

6.2.1. Preprocessing, classification and accuracy of 1996 classification

The final total RMS error in the georectification of 1996 SPOT HRV image was 0,884. This corresponds approximately to 18 meters. Based on visual estimation the rectification was more distorted in the northern parts of the island, where errors were more than 1 pixel.

In 1996 approximately 8% of the landscape was non-vegetated, 46% vegetated and 45% forested (Table 14 & Figure 24). The overall accuracy is 83,0%, which would mean that the overall accuracy for the change detection would be 68,48% (0,825 * 0,83 = 0,68475) (Table 15). Although this is relatively good accuracy for automated change detection analyses, it still contains 31,52% possibility for error. The kappa figure is 0,784, making the classification 78,4% better than one created on completely random basis.

Based on producer's accuracy *barren* (80,6%), *semi-open scrubs on grass* (80,8%), *woodland* (92,6%) and *closed forest and scrub* (96,7%) classes included mainly those land covers intended, but high amount of areas that should have been classified as *urban* (66,7%), *low-lying vegetation* (61,1%) or *semi-open scrubs on barren* (72,4%) were misclassified. Considerable amount of *barren* pixels are misclassified as *urban, semi-open scrubs on barren* as *low-lying vegetation* and other vegetated classes as *semi-open scrubs on barren.* From the perspective of user's accuracy *barren, urban, semi-open scrubs on grass, woodlands, closed forests and scrubs* have accuracy figures over 80%, but low-lying *vegetation* (75%) and *semi-open scrubs on barren* (52,5%) have clearly weaker accuracies. *Semi-open scrubs on barren* are systematically mixed up with *low-lying vegetation*, but errors in *low-lying vegetation* were caused by multiple different classes.

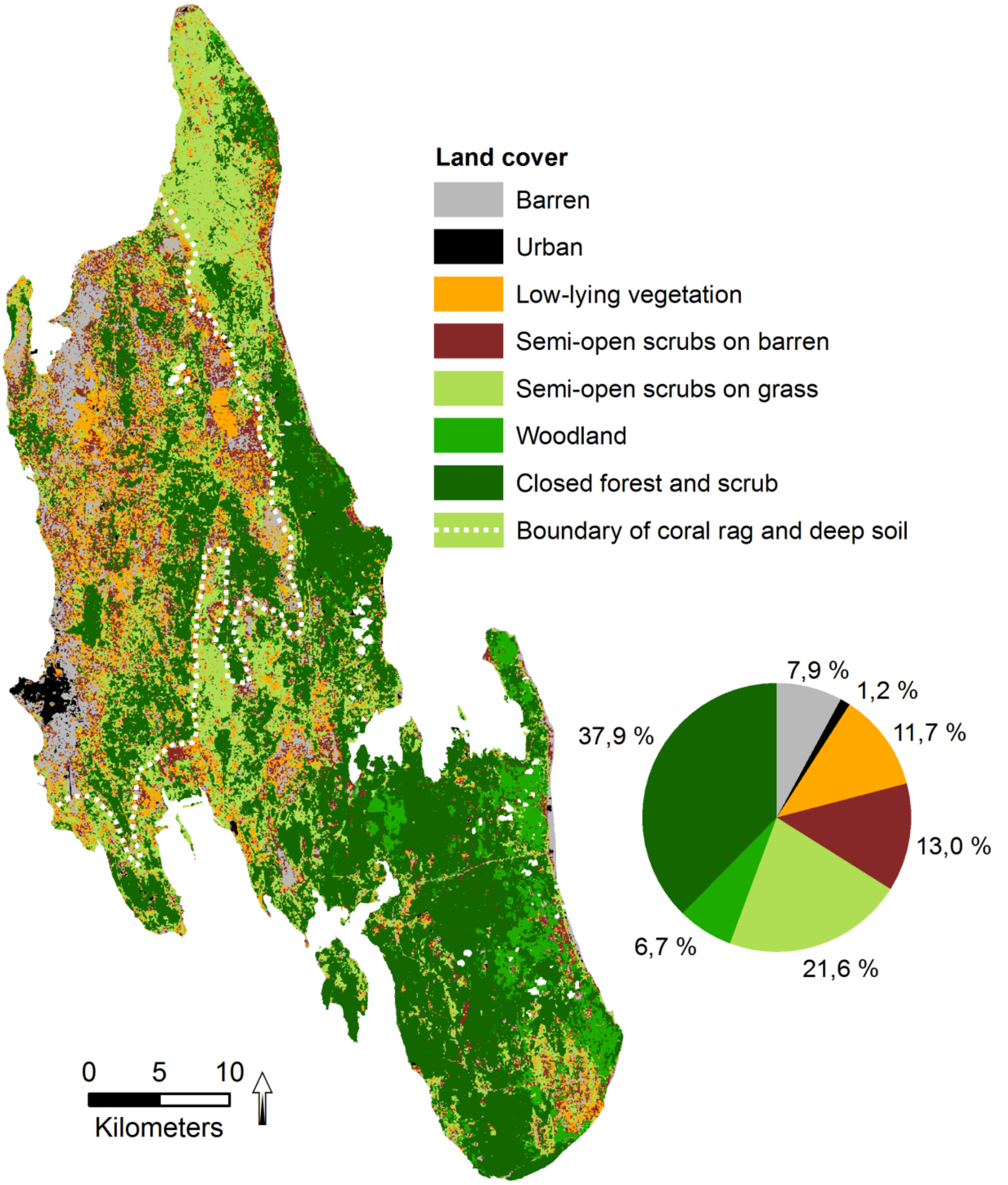


Figure 24. 1996 land cover classification with 7 classes classified from SPOT-3 HRV image dated 30.06.1996.

Table 14. Land cover statistics of Unguja in 1996 based on classification of SPOT-3 HRV (30.06.1996) image.

1996 land cover classification			
	Pixels	Area (km^2)	Percentage of landscape
Barren	132122	118,9	7,9
Urban	19344	17,4	1,2
Low-lying vegetation	195578	176,0	11,7
Semi-open scrubs on barren	216411	194,8	13,0
Semi-open scrubs on grass	360256	324,2	21,6
Woodland	111800	100,6	6,7
Closed forest/scrub	630656	567,6	37,9
Total	1666167	1499,6	100,0

Table 15. Error matrix with accuracy assessment figures for classification of SPOT-3 HRV image from 30.06.1996. Aerial photographs of DoSUP from 1989 were used as reference data.

		Reference data								
	Class	**B**	**U**	**LV**	**SB**	**SG**	**W**	**CF**	**Total**	**User's accuracy**
Classified data	**B**	29	4	1	1	1			36	80,6
	U	1	8						9	88,9
	LV	4		33	3	4			44	75,0
	SB	1		16	21		1	1	40	52,5
	SG	1		2	3	63		2	71	88,7
	W			2			25	1	28	89,3
	CF				1	10	1	118	130	90,8
	Total	36	12	54	29	78	27	122	358	
	Producer's accuracy	80,6	66,7	61,1	72,4	80,8	92,6	96,7		
	Overall accuracy	83,0								
	Kappa	0,784								

B = Barren, U = Urban, LV = Low-lying vegetation, SB = Semi-open scrubs on barren, SG = Semi-open scrubs on grass, W = Woodland, CF = Closed forests and scrubs.

6.2.2. Change at landscape level

At landscape level there has been massive amount of spatial and absolute changes from 1996 to 2009. Only 42,55% of all pixels have not changed, while 57,45% have done so (Table 16). From the changed area 46,77% are caused by swapping and only 10,69% by actual changes between classes. Only in classes *urban* (64,1%) and *closed forests and scrubs* (58,95%) over half of the pixels stayed unchanged. Majority of changes happened between classes of similar spectral and vegetation properties and when similar classes are combined to *barren, vegetated* and *forested* categories in the abrupt classification, changes diminish (Table 17). The effect of combining classes influences percentages of spatial and absolute changes in same proportion and after it the figures are 34,30% for total change, 27,88% for swap, 65,71% for persistence and 12,84% for total absolute value of net change (Table 18). Also the pixel level persistence is over 50% in every class after combining classes.

In general forested land covers have diminished in area and 6,42% of the total landscape are areas where forested classes had changed to other classes. Statistically this area ceded by forests is evenly divided between barren (3,22%) and vegetated (3,2%) land covers (Table 19). Though spatially deforested areas are mainly occupied by other vegetated land covers, while portion of vegetated areas of 1996 are transferred to barren covers. Vegetated class gained 5,48% of total landscape area (13,99% - 8,51% = 5,48%) from forests, but simultaneously lost 2,28% of area (6,04% – 3,76% = 2,28%) to barren classes and barren received rest 0,94% (3,22% - 2,28%) of its net gain directly from forest classes. From classification with seven classes it can be seen that *closed forest and scrubs* (8% of total landscape) and *barren* (1,68%) lost most area, while *urban* (4,9%), *semi-open scrubs on barren* (4,21%) and *woodlands* (1,58%) gained the most. From these only *closed forests and scrubs* (5,8%) lost its area to classes outside its own abrupt classification category (forested), while *urban* (3,14%) and *semi-open scrubs on barren* (1,91%) gained significantly from other categories.

Table 16. Landscape change statistics between 1996 and 2009 at the level of 7 land cover classes.

	Gain	Loss	Persistence	Total change	Swap	Absolute value of net change
Barren	4,07	5,75	2,17	9,82	8,14	1,68
Urban	5,32	0,42	0,75	5,74	0,84	4,9
Low-lying vegetation	8,67	9,22	2,52	17,89	17,34	0,55
Semi-open scrubs on barren	13,56	9,35	3,64	22,91	18,7	4,21
Semi-open scrubs on grass	12,58	13,04	8,58	25,62	25,16	0,46
Woodland	5,72	4,14	2,57	9,86	8,28	1,58
Closed forest/scrub	7,54	15,54	22,32	23,08	15,08	8
Total	57,46	57,46	42,55	57,46	46,77	21,38

Table 17. Change matrix between 1996 and 2009 at the level of 7 land cover classes

		2009							
		B	U	LV	S-O B	S-O G	W	C F/S	Total 1996
1996	B	2,17	1,76	1,04	1,72	0,79	0,14	0,3	7,92
	U	0,12	0,75	0,06	0,1	0,05	0,02	0,07	1,17
	LV	1,22	0,91	2,52	2,96	2,87	0,31	0,95	11,74
	S-O B	1,45	1,13	1,94	3,64	2,36	0,94	1,53	12,99
	S-O G	0,64	0,69	3,29	3,64	8,58	1,1	3,68	21,62
	W	0,19	0,22	0,43	1,55	0,74	2,57	1,01	6,71
	C F/S	0,45	0,61	1,91	3,59	5,77	3,21	22,32	37,86
	Total 2009	6,24	6,07	11,19	17,2	21,16	8,29	29,86	100

B = Barren, U = Urban, LV = Low-lying vegetation, S-O B = Semi-open scrubs on barren, S-O G = Semi-open scrubs on grass, W = Woodland, C F/S =Closed forest scrub

Table 18. Landscape change statistics between 1996 and 2009 at the level of 3 land cover classes (abrupt classification).

	Gain	Loss	Persistence	Total change	Swap	Absolute value of net change
Barren	7,51	4,29	4,8	11,8	8,58	3,22
Vegetated	17,75	14,55	31,8	32,3	29,1	3,2
Forested	9,04	15,46	29,11	24,5	18,08	6,42
Total	34,3	34,3	65,71	34,3	27,88	12,84

Table 19. Change matrix between 1996 and 2009 at the level of 3 land cover classes (abrupt classification).

		2009			
		Barren	Vegetated	Forested	Total 1996
1996	Barren	4,8	3,76	0,53	9,09
	Vegetated	6,04	31,8	8,51	46,35
	Forested	1,47	13,99	29,11	44,57
	Total 2009	12,31	49,55	38,15	100

6.2.3. Rates of forest change

From the abrupt classifications transition matrix and the change matrix of all seven classes it can be seen that 24,9 - 29,11% (373,2 - 436,5 km^2) of the total landscape is covered by persistent forests, depending on if forests changed from closed to more open woodlands are considered stable or not. These stable or persistent forests covered 66,05 - 77,26% of 2009 and 56,1 - 65,62% of 1996 forests and are located inside Jozani, Kiwengwa-Pongwe and Masingini government forest areas, just outside these protection areas and in the agroforests of deep soil region, but the largest concentrations are the community forests in South district. Persistent forests are a rarity near to the coastline and existed only between Kizimkazi and Jozani and in Uzi Island. When the amount of persistent forests are proportioned against all forests existing either in 1996, 2009 or in both years their share is 46,4 - 54,30% and the rest 45,70 – 53,58% (367,40 – 430,70 km^2) are forests that vanished or generated during the time period. Evidently this points to high swapping of forests and actually 73,80% of forest changes are caused by swapping and only 26,20% by permanent loses from forested to barren or vegetated. Both absolute and swapping changes are mainly caused by transitions from forested classes to *semi-open scrubs on grass, semi-open scrubs on barren* and to *low-lying vegetation,* in this order. From forested classes *closed forests and scrubs* cause more net and swapping changes obviously because of its bigger size. Though when proportioned to the total changes of the class, within *woodlands* swapping cause 83,98% of changes, while in *closed forests and scrubs* this is only 65,34%.

When looking at the transitions based on the change classes, land cover changes classified as *deforestation* cover 134,1 km^2 in the gradual classification and 231,7 km^2 in the abrupt one. These figures are 9,0% and 15,5% of the total landscape and 20,2% and 34,9% from the total forest cover in 1996. If the deforestation would be evenly divided between the 13 years of the study period, it would mean that 10,32–17.82 km^2 of forests have disappeared annually (Figure 25) (Table 20). If reforestation is not taken into account and it is assumed that the absolute amount of deforestation stays the same, it would mean that all the remaining 565 km^2 of forests, scrubs and woodlands would go through deforestation in 32 to 55 years and all the remaining stable forests in 21 to 42 years.

Table 20. Statistical figures for abrupt and gradual change classification.

	Gradual classification		**Abrupt classification**	
Change class	Area (km^2)	Area from total landscape (%)	Area (km^2)	Area from total landscape (%)
Reforestation	63,7	4,3	135,5	9,0
Forest improvement	87	5,8	N/A	N/A
Stable forest	373,2	24,9	436,5	29,1
Forest degradation	145,8	9,7	N/A	N/A
Deforestation	134,1	8,9	231,7	15,5
Revegetation	56,5	3,8	56,5	3,8
Vegetation imrpovement	78,4	5,2	N/A	N/A
Stable vegetation	294,6	19,6	476,9	31,8
Vegetation degradation	103,9	6,9	N/A	N/A
Devegetation	90,5	6	90,5	6
Stable non-vegetated	71,9	4,8	71,9	4,8
Total	1499,6	100	1499,6	100

Only 9,48 – 16,38% of the happened deforestation is caused by transitions from forested surfaces to barren or urban land covers. These changes concentrate to coastal areas, government forests and close to Zanzibar Town, while smaller patches spread all around coral rag region. Majority of deforestation is changes from forests to other vegetation (83,62 – 90,52%). The biggest individual changes are from forested classes to *semi-open scrubs on grass* and the 6,51 percentage point range in total deforestation estimates came from different approaches to this class. Changes from forested classes to vegetated ones spread quite steadily all around the island, but some larger concentrations are found from the outskirts of Zanzibar Town, Fumba peninsula, Uzi island, coastal areas of Muyuni, Makunduchi and Kizimkazi, tip of Michamvi peninsula, north from the tip of Kiwengwa-Pongwe forest reserve and from coastal region between Matemwe and Nungwi.

The gradual classification class f*orest degradation* cover 145,8 km^2, 9,7% from the total landscape and 21,92% from the forests of 1996. If it is assumed that the absolute amount of degradation has been stable during the study period, this would mean that

11,2 km^2 of forests would have been degradading annually. With this continuous pace all the remaining forests and scrubs would degrade in 50 years if forest improvement and reforestation is not taken into account. Majority (66,98%) of forest degradation is caused by changes from forested classes to *semi-open scrubs on grass* and only 33,02% is caused by transitions from *closed forests and scrubs* to *woodlands.* The former is part of deforestation in abrupt classification and its spatial distribution is mentioned in that section. The latter appears mainly in the southern peninsula of Unguja where majority of *woodlands* are located. Also some of the agroforest areas especially in Kidanzini ward have gone through changes from dense to more sparse.

63,7 to 135,5 km^2 of Unguja *reforested* between 1996 and 2009. This means that 7,63 - 16,24% of the non-forest areas of 1996 forested during the study period and 11,27–23,98% from the total forests existing in 2009 are created by reforestation. If the process is equally divided between the study years, this means that 4,9–10,4 km^2 of new forests is created annually. Transitions from barren surfaces to forests cover 7,89 km^2, which is 0,53% from the total landscape and 5,82–12,39% from the class, without any major spatial concentrations. Rest 55,81–127,61 km^2, 3,77–8,47% from the landscape and 87,61 – 94,18% from the class are changes from vegetated to forested. These concentrated to government forest areas of Jozani, Kiwengwa-Pongwe, Kibele and Dunga, to all major agroforests and to the area between Jozani and Kiwengwa-Pongwe government forests.

Forest improvements cover 87 km^2, 5,8% from the total landscape and 15,40% from 2009 forests and woodlands. This would be 6,69 km^2 when divided steadily between the 13 study years. Changes from *semi-open scrubs on grass* to *closed forests and scrubs* or *woodlands* contribute 71,78 km^2, 4,78% from total landscape and 82,51% from the class and the rest 15,19 km^2, 1,01% from total landscape and 17,49% from the class are changes from *woodlands* to *closed forests and scrubs.* The former concentrates to government forests and to the area between Jozani and Kiwengwa-Pongwe as mentioned earlier, while the patches of the latter are so small that no clear concentrations exist.

If reforestation is subtracted from deforestation to calculate *net forest change* or more precisely *net deforestation* in this case of negative trend, it is 70,4–96,2 km^2 during the 13 year period and 5,42–7,4 km^2 annually if divided even between the years. This is 10,6–14,5% from the total forest cover in 1996. If the absolute amount of forest decline would stay unchanged, it would mean that the 572 km^2 of forest cover still remaining in 1996 would disappear completely in 76 to 104 years. Combined *deforestation* and *forest degradation* covers 279,8 km^2, which would mean annual combined deforestation and degradation pace of 21,52 km^2 if divided evenly. If the process would happen only in the remaining gradual classification stable forests with the same pace they would be wiped out in 17 years and 4 months, but with combined reforestation and forest improvement speed of 150,7 km^2 during the 13 years and 11,59 km^2 annually, 197,07 km^2 of forests and scrubs would have grown back during this process. When reforestation and forest improvement figures are subtracted from deforestation and degradation *the final net forest deforestation/degradation* figure is 129,2 km^2 and 19,42% from the total landscape and 9,94 km^2 and 19,33% from the forest stock of 1996. This would mean that all of the remaining 572 km^2 of woodlands and closed forests and scrubs would go through degradation or complete deforestation in fore coming 57 years if the absolute amount of change would stay equal. This figure also includes changes from *closed forest/scrub* to *woodlands.*

Based on the FAOs (1995) equations the annual rate of forest change is -1,1886 when calculated from the abrupt classification and -0,8247 when calculated from gradual classification. These relative and the two absolute figures provide significantly different future estimations (Figure 26). The annual forest cover change rate drawn from the gradual classification would argue that forest cover drops below 500 km^2 only after 2025 and below 400 km^2 somewhere in 2050s, while the abrupt classification change rates would argue that these declines happen around 2020 and 2040. However if the absolute decline continues to be 7,4 km^2 annually the forest cover is below 400 km^2 already in 2030s and below 300 km^2 in 2040s.

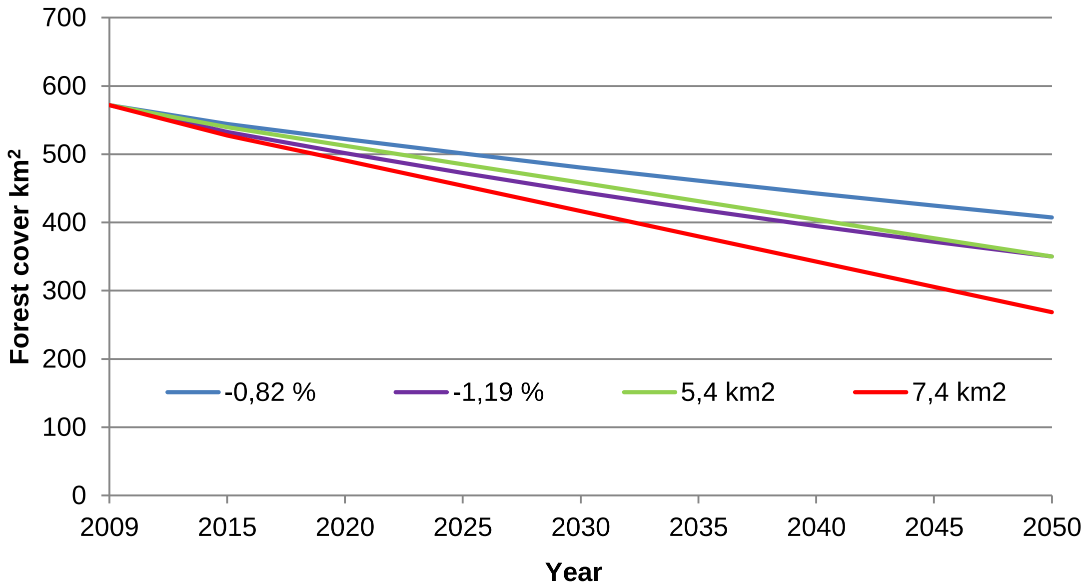


Figure 26. The forest cover scenarios until 2050 based on annual forest change rates and the absolute areas of annual forest cover decline.

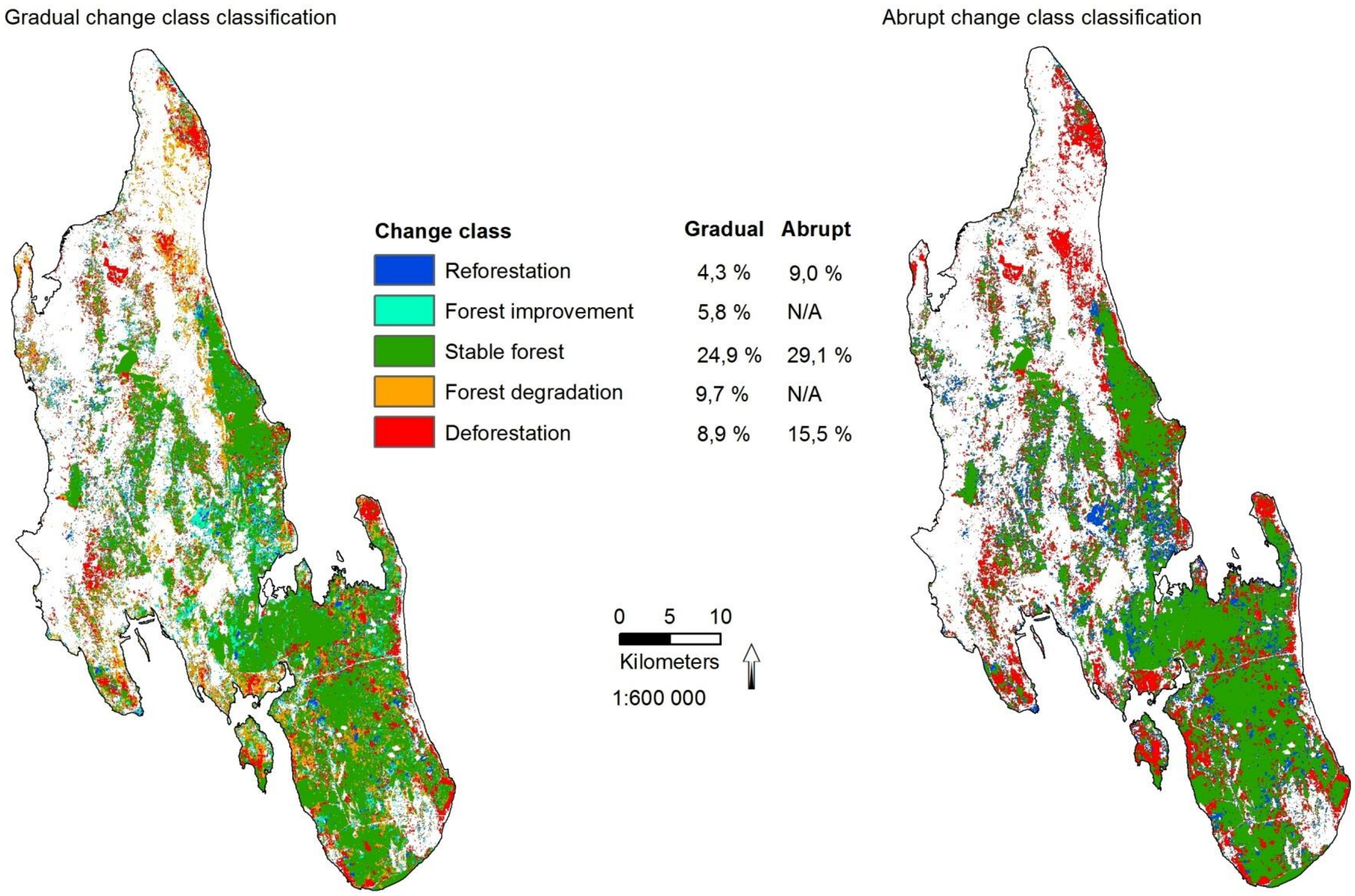


Figure 25. Forest changes of Unguja between 1996 and 2009. Gradual classification deals semi-open scrubs on grass as vegetated class, while in the gradual classification the class is considered as degraded forest class.

6.2.4 The key areas of forest change

The z-score of Moran's I Spatial Autocorrelation calculations for the test site peaks at 1000 m cut off distance and its Moran's I index at this threshold is 0,46 indicating relatively high clustering (Table 22).

Table 22. Outcomes of Moran's I calculations for the chosen test site

Moran's I			
Threshold (m)	Moran's I Index	z-score	p-value
500	0,62	380	0,0000
750	0,54	428	0,0000
1000	0,46	441	0,0000
1250	0,40	432	0,0000
1500	0,34	410	0,0000

The key areas or hotspots of deforestation form altogether over 40 patches, wherefrom 9 are larger than 3,5 km^2 by their abrupt KDE borders (Figure 27). Only two of these are in deep soil, one within Chaani government forestry site and one at the outskirts of Zanzibar Town. Integrating factor between these two hotspots, one larger in Michamwi peninsula and the smaller ones close Kizimkazi, Makunduchi and Paje is that their borders are almost as large in the grdual classification KDE as they are in the one based on the abrupt classification. This indicates that in these areas changes have been from forest to completely different land covers, possibly caused by clear cuts, permanent agriculture and urban expansion. The six remaining large hotspots in Uzi Island, Fumba peninsula, Muyuni, Jozani, Kiwengwa/Matemwe and Matemwe/Nungwi all have abrupt classification KDE borders surrounding the gradual ones or they are not even considered as hotspots by gradual classification standards. This indicates that they are formed by a core area of absolute land cover changes surrounded by degradation of forests or that they are created by degradation only. Also the smaller hotspots concentrate on the coral rag side of the island and they are most present in the South district.

There are altogether 14 reforestation hotspots and these are considerably smaller than the key deforestation areas (<3,5 km^2). The gradual classification KDE did not recognize any reforestation hotspots, and even when forest degradation and improvement were included, only one significant reforestation hotspot was found within Kibele government forestry area. In abrupt classification KDE hotspots were recognized within Jendele and Jozani-Chawka Bay forest areas and north from Jozani-Chwaka Bay.

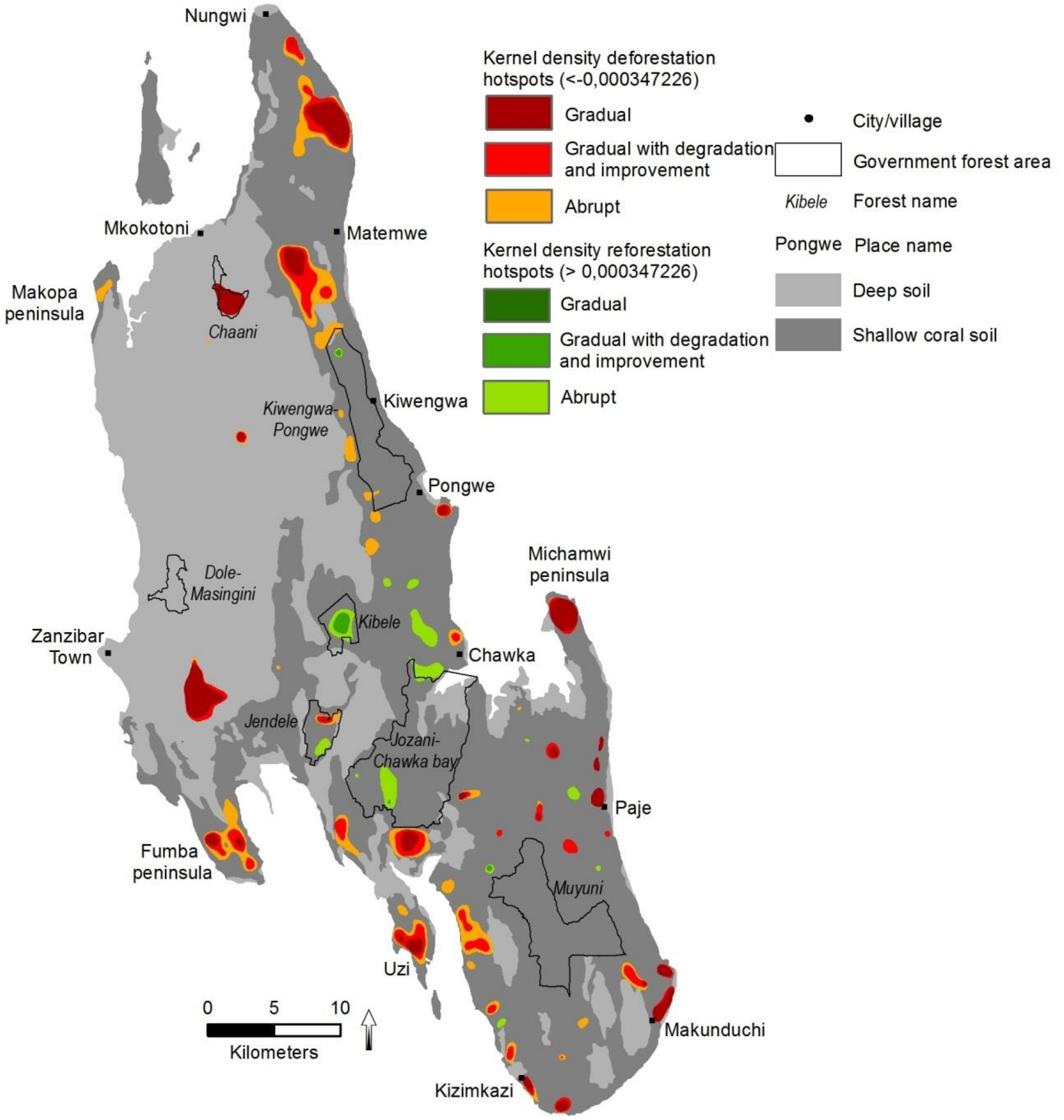


Figure 27. The key areas or hotspots of deforestation and reforestation calculated with Kernel Density Estimation with 1 km cut off distance from the gradual and abrupt classification classes "deforestation" and "reforestation" and also by using gradual classification classes "forest degradation" and "forest improvement". The outcomes are laid on the landscape/soil delineation and government forest areas are included.

6.3. Influence of environmental factors on forest changes

6.3.1. Change differences between forest types

Unguja's forests can be divided to communal *indigenous forests* (59,6–62,9% of total forest cover), *agroforests* (21,8–27,6%), *government protected forests* (10,3–15,1%) and to *government forestry forests* (0,6–1,6%) (Table 23 & Figure 28). Forests cover 73,88–85,80% of total land area in protected areas, while within other entities this share is lower and as low as 12.76–38,04% within *government forestry forests.* The proportion of *indigenous forests* from all forests have gone from 62,94% to 59,61%

between 1996 and 2009. These lost 3 percentage points have been gained by *government protected forests,* while other forests categories have stayed relatively stable*.* However this does not mean that the absolute area lost by *indigenous forests* is gained by the *government protected forests*, but rather the total amount of forests have dropped drastically, while the cover is increased slightly within the protection zones. In absolute square kilometers *indigenous communal forests* have diminished significantly in size (79,65 km^2), *agroforests* slightly (21,59 km^2), *government forestry forests* negligibly (0,57 km^2) and the *government protected forests* have gained area (5,61 km^2). Based on the FAO (1995) function the annual forest cover change is -1,60% in *indigenous communal forests,* -1,05% in *agroforests,* -0,58 in *government forestry forests* and 0,59 in *government protected forests.* In 2009 86,10% of *government protected forests,* 79,98% of *communal indigenous forests,* 64,67% of *agroforests* and only 36,22% of *government forestry forests* were stable forests, which have not changed between 1996 and 2009. The *government protected forests* have higher share from the stable forests than they have from all forests of 1996 and 2009, while in agroforest and forestry areas this share is lower, which indicates that there is more swapping happening in the last two forest types, while protected forests are more persistent.

Table 23. Forest and class statistics of four different forest types.

Forest category	Indigenous communal forests	Agroforests	Government protected forests	Government forestry forests	Total
Forests (km^2) in 1996	420,62	168,78	71,03	7,81	668,24
Forests (km^2) in 2009	340,97	147,19	76,64	7,24	572,03
Stable forests (km^2)	272,70	95,20	65,99	2,62	436,52
Share from all of the forests in 1996 (%)	62,94	25,26	10,63	1,17	100
Share from all of the forests in 2009 (%)	59,61	25,73	13,40	1,27	100
Share from all of the stable forests (%)	62,47	21,81	15,12	0,60	100
Share from all forests (%)	61,5	27,6	10,3	1,6	100
Share of 1996 forests from the total class area (%)	56,08	26,39	79,52	38,04	44,57
Share of 2009 forests from the total class area (%)	45,46	23,01	85,80	35,27	38,15
Share of stable forests from the total class area (%)	36,36	14,88	73,88	12,76	29,11
Share of the stable forests from all of the 2009 forests within the forest type	79,98	64,67	86,10	36,22	76,31

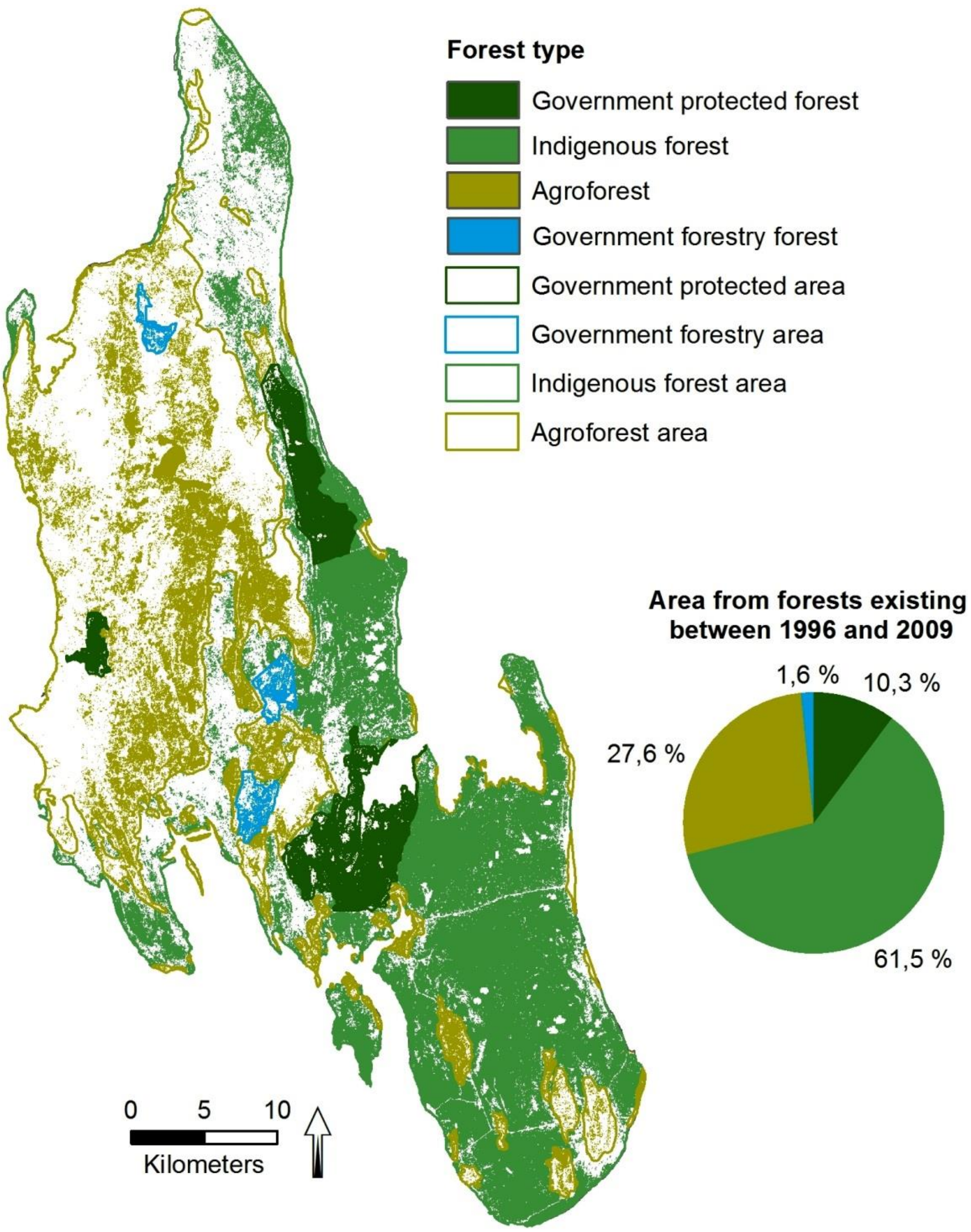


Figure 28. The spatial distribution of the four forest types of Unguja and their share from all the forests ever existed between 1996 and 2009.

Indigenous forests share of forest degradation in the gradual classification is higher than its share generally from the forests (Figure 29). In other words forest degradation happens more frequently in *indigenous forests* than in other forest types. Reforestation and forest improvement on the contrary happen more seldom. Deforestation and amount of stable forest are in line with the general share of forests. Shares of deforestation (36,3%), forest degradation (29,7%), reforestation (40,2%) and forest improvement (32,5%) are significantly higher in *agroforest* than its share range from total forest area, while amount of stable forest (20,1%) is lower, which implies that agroforests are rather dynamic in their nature. The share of deforestation (1,8%) and forest degradation (2,9%) are extremely low in *government protected forests,* though also reforestation (6,9%) and forest improvement (9,5%) levels are lower than the general share of forests, while proportion of stable forests (16,7%) is significantly higher. *Government forestry forests* have higher share from forest improvement (4,4%), deforestation (2,7%) and reforestation (1,6%), while the share from stable forests (0,6%) is lower than the share of general forest cover. Within all other forest types over 90% of changes are from forested to vegetated and vice versa, but within *government forestry* areas changes from forested to barren cover 28% of all losses.

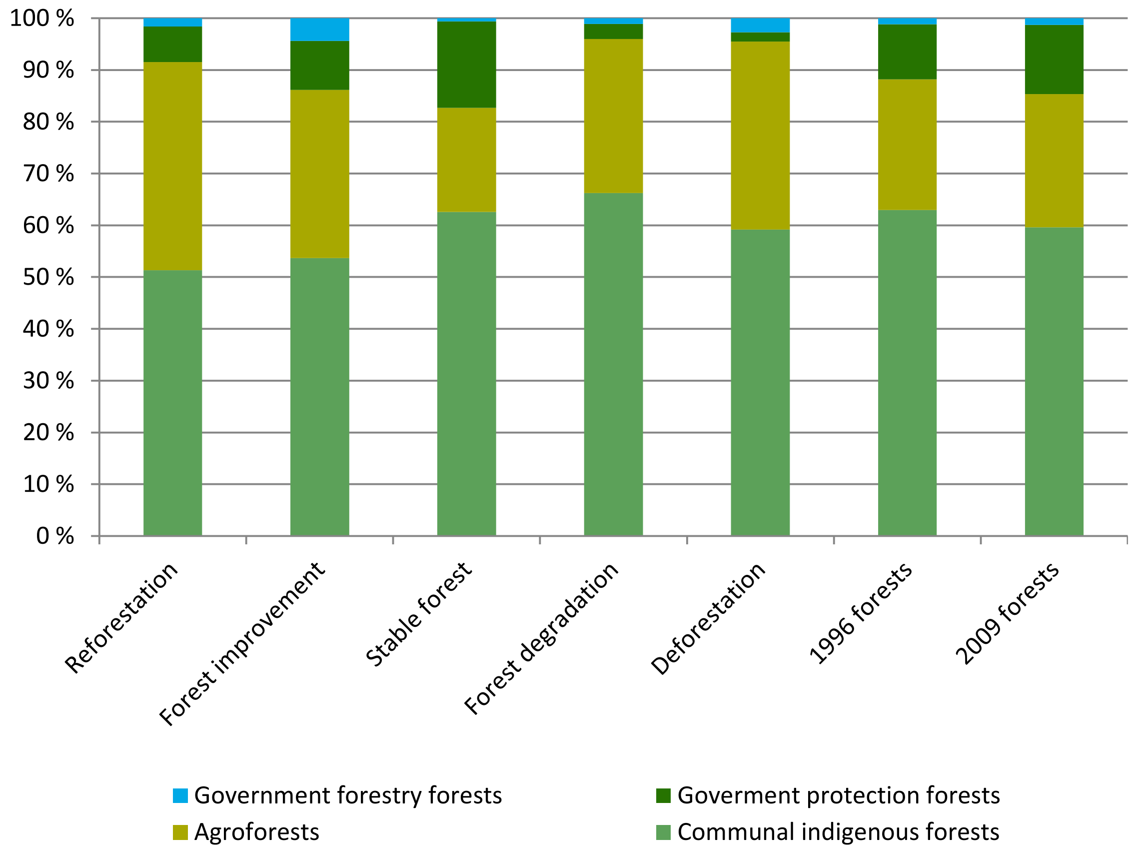


Figure 29. The shares of different forest types from changes, stability and overall forest cover.

When looking at the landscape change statics of each forest type, it is clear that there have been enormous differences in the change processes (Table 24). For example swapping causes 63,19% of forest changes in *indigenous forests,* 64,24% in *government protected areas,* 82,78% in *agroforests* and 94,20% in *government forestry forests.* After the baseline situations of forest cover is acknowledged some generalizations can be done (Table 25). *Indigenous forests* could be described as spatially relatively stable forests facing high net deforestation, but still having generally a higher forest cover, *agroforest* and *government forestry forest* as areas of unstable forests with high locational forest changes and moderate net deforestation and *government protected forests* as areas of extremely high cover and stability of forests with moderate net forest gain.

Table 24. Landscape change statistics of different forest types

Landscape change statistics for each forest type*												
Communal indigenous forests												
	Barren	Vegetated	Forested	Total 1996		Gain	Loss	Persistence	Total change	Swap	Absolutevalue of net change
Barren	0,90	2,14	0,62	3,65	Barren	3,65	2,76	0,90	6,40	5,51	0,89
Vegetated	1,94	29,84	8,49	40,27	Vegetated	20,15	10,43	29,84	30,58	20,86	9,73
Forested	1,71	18,02	36,36	56,08	Forested	9,10	19,72	36,36	28,82	18,21	10,62
Total 2009	4,55	50,00	45,46	100,00	Total	32,91	32,91	67,09	32,91	22,29	21,24
Agroforests											
	Barren	Vegetated	Forested	Total 1996		Gain	Loss	Persistence	Total change	Swap	Absolute value of net change
Barren	10,17	6,19	0,38	16,74	Barren	12,77	6,57	10,17	19,35	13,14	6,20
Vegetated	11,62	37,52	7,74	56,88	Vegetated	16,53	19,36	37,52	35,90	33,07	2,83
Forested	1,16	10,35	14,88	26,39	Forested	8,13	11,50	14,88	19,63	16,25	3,38
Total 2009	22,94	54,05	23,01	100,00	Total	37,44	37,44	62,56	37,44	31,23	12,41
Government protected forests											
	Barren	Vegetated	Forested	Total 1996		Gain	Loss	Persistence	Total change	Swap	Absolute value of net change
Barren	0,07	0,64	0,87	1,58	Barren	0,63	1,51	0,07	2,15	1,26	0,88
Vegetated	0,27	7,58	11,04	18,89	Vegetated	5,92	11,31	7,58	17,23	11,84	5,39
Forested	0,36	5,28	73,89	79,53	Forested	11,91	5,64	73,89	17,56	11,28	6,27
Total 2009	0,70	13,50	85,80	100,00	Total	18,46	18,46	81,54	18,46	12,19	12,54
Government forestry forests											
	Barren	Vegetated	Forested	Total 1996		Gain	Loss	Persistence	Total change	Swap	Absolute value of net change
Barren	0,54	1,44	0,22	2,21	Barren	13,71	1,67	0,54	15,37	3,33	12,04
Vegetated	6,62	30,86	22,27	59,75	Vegetated	19,63	28,89	30,86	48,52	39,26	9,26
Forested	7,08	18,19	12,77	38,04	Forested	22,49	25,27	12,77	47,76	44,99	2,78
Total 2009	14,25	50,49	35,26	100,00	Total	55,83	55,83	44,17	55,83	43,79	24,08

**All the values are percentages of that particular forest type*

Table 23. Change ratios of different forest types

	Unguja	Communal indigenous forest	Communal agroforest	Government protected forests	Government forestry forest
Forest*	1	1,22	0,64	1,71	1,13
Stable forests*	1	1,03	0,79	1,49	0,39
Total change *	1	0,97	1,24	0,42	1,73
Swapping *	1	0,83	1,40	0,37	2,20
Net deforestation *	1	1,36	0,82	-0,57	0,38

** Values are proportioned to the average total change, swapping and net deforestation of of all forest areas existing in 1996, 2009 or in both years*

6.3.2. Forest change connected to distance measurements

Distance from coast influences deforestation so that areas close to it are more prone to deforestation. Deforestation reaches its peak approximately 500 meters away from the coastline when it is 66% higher than averagely and starts to decline steadily until 1,6 km, when it is at the same levels as the average deforestation (Figure 30 A). Forest gain is 20 – 27% below whole island averages until 1,1 km away from coast and after this approximately 10% below normal figures within the remaining 3 kilometers. The amount of stable forests is generally 25 to 30% below averages until 1,5 kilometers from the coast, after which it keeps steadily increasing. When reforestation and deforestation are combined as net deforestation, these figures are over 1 until two kilometers from the shoreline and peak around 500 meters away, when net deforestation is 2,9 times more common than averagely (Figure 30 B). These outcomes suggest quite directly that forest losses happen more frequently in the area until 1,5 - 2 kilometers away from the coastline. Especially serious the situation is between 0,2 – 1,1 kilometers away where the net forest loss is over 2 times higher than normally. The net deforestation trend is rather linear, however there is some nonlinearity and the R^2 value for the linear trend line is only 0,7277. The magnitude and overall importance of deforestation near coast is rather significant, since based on the cumulative percentages of net deforestation 63% of the process happens within 3 kilometers from the coast and 43% in areas where deforestation is most severe (Figure 31).

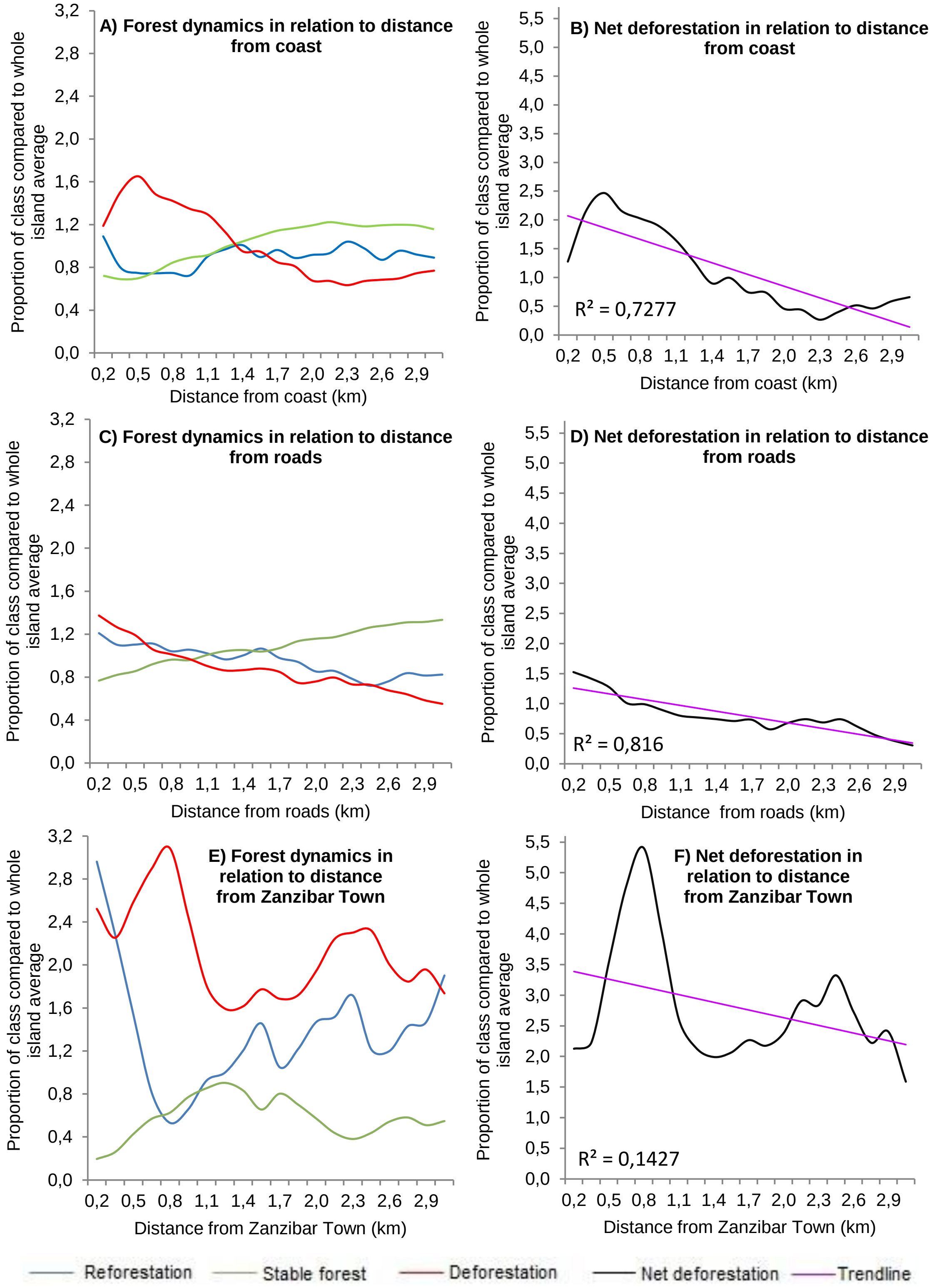


Figure 30. Relation of proportional reforestation, stable forests and deforestation to distance from coast (A), roads (C) and Zanzibar Town (E) and relations of proportional net deforestations and their trendlines to distance from coast (B), roads (D) and Zanzibar Town (F).

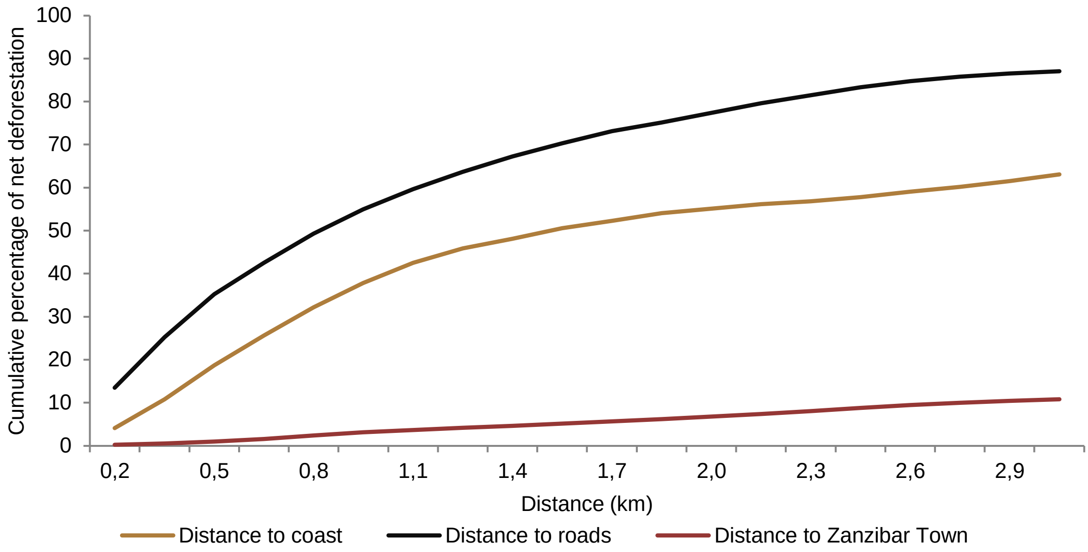


Figure 31. Cumulative percentage of deforestation from different sources.

In Unguja it seems that forests near roads are less prone to deforest than forests in coastal settings or near the capital, but these are still disappearing faster than forest averagely on the island. Forest loss is highest just outside roads (37% more than on average), reaches average levels around kilometer away and keeps declining steadily from here on until 3 kilometers away (Figure 30 C), when it is only 55% from average deforestation. Forest gain works in a similar manner, but having maxima only 21% above average and declining only 19% below average around 3 kilometers. Stable forests increase steadily from 23% less than normally in first zone to 33% more around 3 kilometers. Net deforestation tells similar outcomes and the forest loss is highest right next to roads where it is 52% higher than normally (Figure 30 D). It declines to normal levels after 1,2 kilometers and is below average 2,2 kilometers onwards. The trend is clearly linear and the trendline gets R^2 value of 0,8160. 87% of net deforestation happens within 3 kilometers from roads and 42% of happens within 700 meters from roads, where deforestation is most rapid.

Both deforestation and reforestation were more common in areas just outside Zanzibar Town in 1996, however the reforestation rates start to decline rapidly, while deforestation rates only keep increasing (Figure 30 E). The highest deforestation rates are achieved 800 meters away from the capital when they are 3,1 times higher than on average. The peak of deforestation levels around 1,5 kilometers away, though the process is continuously over 1,5 times more common within the whole three kilometers study zone. The amount of stable forest never achieves average levels within the study zone and varies between 20 – 99% from the average, being highest in the zone between 1 – 1,5 kilometers away where Masingini forest is located. The net deforestation rates show even more staggering figures (Figure 30 F). In the area between 0,5 – 1 kilometers from the city border deforestation was 3,5 times more common than on average and peaks at 800 meters away (5,4). The net deforestation values do not go below 1,5 within the whole study zone and generally the figures are over 2,5. The trend of deforestation is not linear by any means near the capital and linear trendline achieves coefficient of determination of only 0,1427. When interpreting the cumulative percentages of net deforestation it can be seen that in large scale the distance to Zanzibar Town holds only minor importance. From all of the deforestation happening in Unguja only 11% happens within 3 kilometers from the capital and only 4% in the zone where deforestation is most rapid.

6.3.3. Environmental factors explaining spatial patterns of deforestation

The binary logistic regression overall accuracies of single variable models range between 51,5 – 60,5%, while the anti-model estimating all the observations as 0 (non-forested) would reach the accuracy of 51,9% (Table 26). Therefore the best model (distance to coast) is able to estimate only 8,6 percentage points better than the anti-model, however when estimating deforested pixels the accuracy is 61,5 percentage points better.

Table 26. Accuracies, coefficients, Nagelkerke R^2 and Wald statistics of single variable binary logistic regression models explaining the distribution of deforested and stable forest cells.

	Deforested accuracy (%)	Stable accuracy (%)	Overall accuracy (%)	B	Nagelkerke R Square	Wald
Distance to coast*	61,5	59,7	60,5	-0,279	0,117	108,4
Kernel density of buildings*	26,3	89,5	59,1	0,007	0,089	38,7
Mean NDVI*	52,4	64,8	58,9	-20,621	0,041	43,2
Distance to Zanzibar Town*	47,3	63,7	55,8	0,018	0,011	11,7
Government protection status*	16,2	98,3	55,7	-2,461	0,101	59,0
Government forestry status*	6,3	99,6	54,7	2,776	0,043	21,4
Soil**	19,9	81,9	52,1	-0,076	0,001	2,0
Distance to main and secondary roads	0	100	51,9	-0,013	0,000	0,1
Mean elevation**	11	88,9	51,5	0,005	0,002	1,9

***)** *Statistically significant at level 0,05. **) Statistically significant at level 0,5*

Distance to coast is by far the most explanatory variable based on overall and deforested accuracy, Nagelkerke R^2 and Wald statistics, but also kernel density of buildings, mean NDVI and government protection status have high explanatory value. Distance from Zanzibar Town has relatively good deforested and overall accuracies, but its Nagelkerke R^2 and Wald figure are low, while the situation is opposite to government forestry status. Surprisingly the soil, distance to roads and mean elevation are not usable variables for explaining deforestation in Unguja.

The regression coefficient of the best variables show that increasing distance from the coast, increasing mean NDVI and government protection status would reduce the risk of deforestation, while increasing kernel density of buildings, distance from Zanzibar Town and status as government forestry forest would increase the risk of deforestation. The coefficient of "distance to Zanzibar Town" is rather peculiar since average deforestation was over 5 times more common in the areas near the capital than on average and in related literature vicinity of large cities has increased deforestation (Ludeke et al. 1997; Verburg et al. 2004). The deforestation concentrates into the coral rag region and its furthest corners, which then creates this peculiar correlation with distance from the capital, although the matters are not related in reality.

In correlation analyses none of the variables have extremely strong correlations (> 0,7 or < -0,7), which would cause problems of multicollinearity and require dropping variables out from the multivariate regression analyses (Table 27). However some significant correlations (> 0,4 or < -0,4) are found, which should be kept in mind when analyzing the final outcomes.

Table 27. Bivariate Pearsons correlation analysis with two-tailed test of statistical significance showing weak or moderate multicollinearity between the explanatory variables.

	Coast	NDVI	Kernel	Z Town	GPS	GFS	Soil	Roads	Elevation
Coast[a]	1	-,155**	,220**	-,424**	-0,051	,160**	-,578**	-0,028	,599**
NDVI[b]	-,155**	1	-,158**	-0,018	,186**	-,165**	,294**	,211**	-,186**
Kernel[c]	,220**	-,158**	1	-,530**	-0,031	,054*	-,561**	-,159**	,211**
Z Town[d]	-,424**	-0,018	-,530**	1	-,210**	-,127**	,625**	,305**	-,390**
GPS[f]	-0,051	,186**	-0,031	-,210**	1	-,059*	,099**	-0,005	-0,014
GFS[g]	,160**	-,165**	,054*	-,127**	-,059*	1	-,264**	0,013	,346**
Soil[h]	-,578**	,294**	-,561**	,625**	,099**	-,264**	1	,260**	-,658**
Roads[i]	-0,028	,211**	-,159**	,305**	-0,005	0,013	,260**	1	-,064*
Elevation[j]	,599**	-,186**	,211**	-,390**	-0,014	,346**	-,658**	-,064*	1

[a)] Distance to coastline, [b)] Mean NDVI in 1996, [c)] Kernel density of buildings with 6500 m cutoff distance, [d)] Distasnce to Zanzibar Town in 1996, [E)] Government protection status, [F)] Government forestry status[G)] Soil type[H)] Distance to main and secondary roads [I)] Mean elevation.

When the individual explanatory variables are fed into conditional stepwise binary logistic model, the variable "distance from main and secondary roads" is ignored because of its limited capabilities explaining the model. Also the program modified the order of entrance so that “mean NDVI” is left as last and replaced by “mean elevation”. The best multivariate model for the entire island is reached in the step 5, when variables “distance to coast”, “kernel density of buildings”, “government protection status”, “government forestry status” and “mean elevation” were included into the model, in this order (Table 28). The overall accuracy is 77,7% and deforestation accuracy as high as 83,9%, however still 53 (28,3%) stable forest cells are incorrectly classified as deforested (Table 29). In the light of Nagelkerke R^2 (0,441) the model with highest overall accuracy can be seen to have limitations, but introducing new variables only decreases the overall accuracy of the model. Based on the coefficients the relations of individual variables are the same than in the single variable models (Table 30). The Wald statistics implies that the “distance to coast” is by far the most important variable, followed by “kernel density of buildings”, “mean elevation”, “government protection status” and “government forestry status” in this order.

Table 28. The outcomes of conditional stepwise binary logistic model best explaining deforestation and stable forest locations in Unguja.

Step	Variable entered	Deforested accuracy	Stable forest accuracy	Overall accuracy	Nagelkerke R square
1	Distance to coast	61,5	59,7	60,5	0,117
2	Kernel density of buildings	76,7	69,6	73	0,266
3	Mean elevation	80,5	72,5	76,4	0,340
4	Government protection status	83,5	71,3	77,2	0,424
5	Government forestry status	83,9	71,7	77,7	0,441
6	Distance to Zanzibar Town	81,4	72,3	76,7	0,457
7	Soil type	82,2	71,8	76,8	0,464
8	Mean NDVI	79,9	74,2	77,0	0,474

Table 29. The accuracy and correctly estimated cells of the model best explaining the spatial distribution of deforestation and stable forest in Unguja.

		Predicted			
		Stable forest	Deforested	Total	Correct (%)
Observed	Stable forest	134	53	187	71,7
	Deforested	30	156	186	83,9
	Total	164	209	373	77,7

Table 30. The regression coefficients and Wald statistics of the model best explaining the spatial distribution of deforestation and stable forest in Unguja.

Variable	B	S.E.	Wald	df	Siq.	Exp (B)
Distance to coast	-,676	,046	213,346	1	,000	,509
Kernel density of buildings	,009	,001	66,947	1	,000	1,009
Government protection status	-3,548	,507	48,914	1	,000	,029
Government forestry status	2,580	,652	15,665	1	,000	13,203
Mean elevation	,057	,007	64,945	1	,000	1,058
Constant	0,674	,103	42,410	1	,000	1,962

There neither is severe multicollinearity between the variables in the coral rag region (Appendix 5). When the individual explanatory variables are fed into conditional stepwise binary logistic model for the coral rag region, the variables “distance to main and secondary roads” and “government forestry status” are ignored because they could not explain variations or be statistically significant. The model reached highest overall accuracy (76,7%) when all of the six remaining statistically significant variables are included (Table 31 & 32). However highest the deforestated accuracy (83,4%) is achieved already when the three first variables are introduced. Although the accuracies are rather good, the Nagelkerke R^2 figures imply that there are severe limitations in the model.

The directions of the regression coefficients are the same than for the whole island level modeling, however their order of influence based on Wald statistics varies slightly (Table 33). “Distance to coast” is still by far the most explanatory variable, but “government protection status” is the second most influential, only then followed by “mean elevation”, “kernel density of buildings” and “mean NDVI”. However when the outcomes of the model best explaining deforested cells are looked the “kernel density of buildings” is more important factor than “government protection status” (Table 34 & 35).

Table 31. The outcomes of conditional stepwise binary logistic model best explaining both deforestation and stable forest locations in coral rag region.

Step	Variable entered	Deforested accuracy	Stable forest accuracy	Overall accuracy	Nagelkerke R square
1	Distance to coast	72,2	62,3	67,3	0,226
2	Kernel density of buildings	82,6	68,1	75,4	0,325
3	Government protection status	83,4	68,2	75,7	0,389
4	Mean elevation	83,0	67,7	75,4	0,404
5	Mean NDVI	82,6	70,2	76,4	0,417
6	Soil type	82,8	70,6	76,7	0,421

Table 32. The accuracy and correctly estimated cells of the model best explaining the spatial distribution of both deforestation and stable forest in coral rag region.

		Predicted			
		Stable forest	Deforested	Total	Correct (%)
Observed	Stable forest	107	44	151	70,6
	Deforested	25	120	145	82,8
	Total	132	164	296	76,7

Table 33. The regression coefficients and Wald statistics of the model best explaining the spatial distribution of both deforestation and stable forest in coral rag region.

Variable	B	S.E.	Wald	df	Siq.	Exp (B)
Distance to coast	-,680	,058	135,728	1	,000	,506
Kernel density of buildings	,023	,005	18,242	1	,000	1,023
Government protection status	-2,493	,405	37,860	1	,000	,083
Mean elevation	,060	,014	19,368	1	,000	1,062
Mean NDVI	-18,419	4,742	15,086	1	,000	,000
Soil type	1,484	,751	3,907	1	,048	4,410
Constant	20,797	6,058	11,785	1	,001	1076875765,585

Table 34. The accuracy and correctly estimated cells of the model best explaining the spatial distribution of deforestation in coral rag region.

		Predicted			
		Stable forest	Deforested	Total	Correct (%)
Observed	Stable forest	103	48	151	68,2
	Deforested	24	121	145	83,4
	Total	127	169	296	75,7

Table 35. The regression coefficients and Wald statistics of the model best explaining the spatial distribution of deforestation in coral rag region.

Variable	B	S.E.	Wald	df	Siq.	Exp (B)
Distance to coast	-,587	,052	128,118	1	,000	,556
Kernel density of buildings	,036	,005	56,254	1	,000	1,036
Government protection status	-2,398	,385	38,794	1	,000	,091
Constant	,635	,157	16,245	1	,000	1,886

The multicollinearity of the variables in the deep soil region is also limited and no correlations over 0,7 or under -0,7 are found (Appendix 6). Variable “mean NDVI” is excluded by the model and the order modified significantly. The best modeling outcome is achieved by using the eight other variables. Also the variable “distance from Zanzibar Town” is removed, because it made the same peculiar assumptions that the increasing distance from the capital would increase the deforestation, although it is evident that the deforestation in deep soil region is most severe in the vicinity of the capital. After the re-run the best model was the one including all of the remaining variables (Table 36). This model has overall accuracy of 85,7%, deforested accuracy of

85,4% and Nagelkerke R^2 of 0,668, which all argue that the model functions quite nicely for the deep soil region (Table 37). The direction of influence implied by the variable coefficients are similar than on the whole island level, except also the soil type influences deforestation so that the enclaves of coralline soils are more prone to deforestation (table 36). The influences of different variables are on the other hand highly different than for the whole island or coral rag region. The "distance to coast" is not anymore the most influential variable, but only as explanatory as "government protection status" or "mean elevation". "The kernel density of buildings" however is here the most influential factor. Also the differences in Wald statistics are smaller, arguing that it is not a single factor, but combination of then that models deforestation the best in the western parts of the island.

Table 36. The outcomes of conditional stepwise binary logistic model best explaining deforestation and stable forest locations in deep soil region.

Step	Variable entered	Deforested accuracy	Stable forest accuracy	Overall accuracy	Nagelkerke R square
1	Distance to coast	37,0	89,7	64,9	0,209
2	Kernel density of buildings	60,1	90,3	76,1	0,383
3	Government protection status	78,6	81,5	80,2	0,480
4	Government forestry status	79,2	85,1	82,3	0,529
5	Soil type	80,3	87,2	84,0	0,566
6	Mean elevation	85,4	86,1	85,7	0,668

Table 37. The accuracy and correctly estimated cells of the model best explaining the spatial distribution of both deforestation and stable forest in deep soil region.

		Predicted			
		Stable forest	Deforested	Total	Correct (%)
Observed	Stable forest	31	5	36	86,1
	Deforested	6	35	41	85,4
	Total	37	40	77	85,7

Table 38. The regression coefficients and Wald statistics of the model best explaining the spatial distribution of both deforestation and stable forest in deep soil region.

Variable	B	S.E.	Wald	df	Siq.	Exp (B)
Distance to coast	-,609	,100	37,050	1	,000	,544
Kernel density of buildings	,016	,002	52,383	1	,000	1,016
Government protection status	-10,665	1,724	38,251	1	,000	,000
Government forestry status	2,220	,596	13,860	1	,000	9,209
Soil type	2,328	,530	19,305	1	,000	10,257
Mean elevation	,087	,014	36,953	1	,000	1,091
Constant	-4,615	1,050	19,322	1	,000	,010

Before the deforestation probability mapping the regression model for deep soil region is modified, because the outcomes of mean elevation are not considered reliable and it is seen implausible to predict the future of *government forestry forests.* New correlation coefficients are calculated for the deep soil region without these explanatory variables (Table 39).

Table 39. The regression coefficients used for predicting the deforestation probability in coral rag and deep soil regions

Variable	Coefficients for coral rag model	Coefficients for deep soil model
Distance to coast	-0,587	-0,235
Kernel density of buildings	0,036	0,009
Government protection status	-2,398	-4,149
Government forestry status	n/a	n/a
Soil type	n/a	1,293
Mean elevation	n/a	n/a
Constant	0,635	-1,658

The deforestation probability mapping points out those forest areas that are at the highest deforestation risk, based on the previous regression analyses (Figure 32). The pressure is highest in coastal forests and in agroforest areas near Zanzibar Town. Although the coastal forests are generally facing high risk of deforestation, the probabilities are lower in certain areas because of the low population pressure. These less threaten coastal forests concentrate in to the south-western corner of South district and in to Michamwi peninsula. The government protected areas, interioirs of large indigenous scrublands in the south and the inland agroforest areas are the ones with the lowest probability for deforestation. However the population pressure is so severe in the south-western corner of Masingini forest that it achieves rather high deforestation probability values even though it is officially protected.

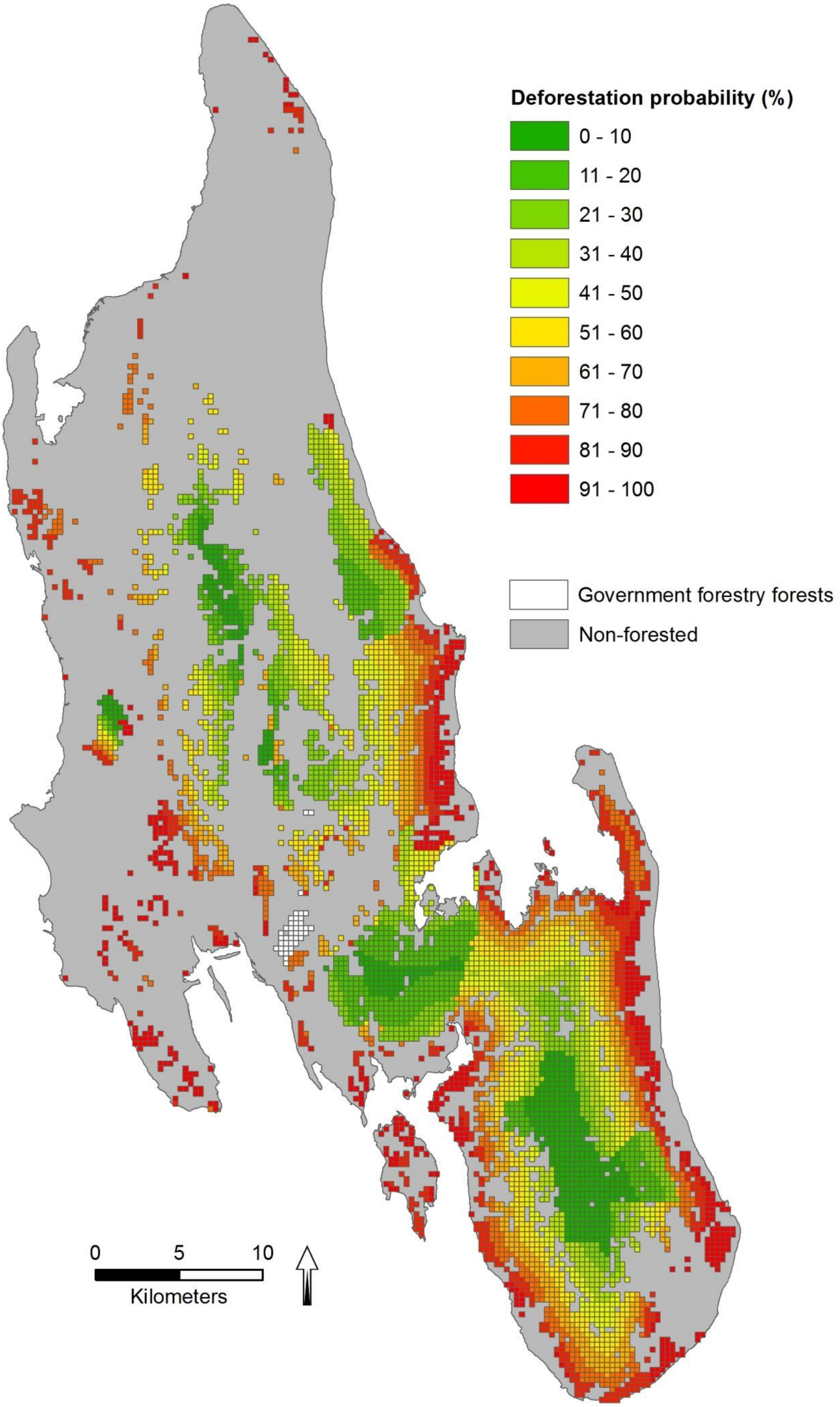


Figure 32. Deforestation probability of the cells still forested in 2009 based on relationship between deforestation and environmental variables between 1996 and 2009.

7. Discussion

7.1. Forests and forest changes in Unguja

7.1.1. Quantity and spatial pattern of today's forest cover

The created classification is the most accurate spatially comprehensive forest coverage estimation available for Unguja, since other similar classifications are either outdated and overestimating the coverage of coral rag scrubs (Woody Biomass Inventory of Zanzibar Islands, 1997) or generalizing the whole coral rag region to a single category of "scrubs" (DoSUP Land Use classification 2009). Based on the classification the forests, scrubs, woodlands and thickets cover approximately 38% of Unguja. However, if this is considered a lot or a little, is a subjective question. Based on FAO (2010) statistics the figure is similar than for Tanzania. However, it is less than in the large continental African countries at the same latitudes (Angola, Congo, Democratic Republic of Congo, Mozambique, Zambia) and more than in the small but highly populated countries of Uganda, Burundi and Rwanda. There is more forests in Unguja than in the other small East African island nations with similar land area and population (Mauritius, Comoros and Reunion) or as in the Caribbean and Oceania islands with similar population densities (300 – 500km^2). As a generalization it could be said that Unguja has relatively a lot of forest cover for an island with extremely high population density and the situation is especially good in the context East African island nations. Though, the different description of forests and methods measuring them makes it extremely problematic to compare the FAO forest statistics (Olander 2008). Therefore the given generalization about the forest cover is rather suggestive than absolute.

The spatial structure of current forests is governed by soil and land use related policies. The coral rag area is significantly more forested (45%) than the deep soil region (23%), even if the government forests located mainly in the coral rag region would not be included into the calculations. The fertile deep soils have been cleared for agriculture and tree plantations already decades and centuries ago, while the low agricultural fertility of the shallow coralline soils have prevented similar actions from happening on the other side of the island (Hettige 1990: 95–98; Krain 1998; Klein & Käyhkö 2008). Though, it is not only the soil fertility that matters. It has been pointed out that the coral soils are more fertile than the deep soils, but they are too shallow, concentrated to potholes, mixed with stones and rugged by coralline outcrops, all eventually making cultivation extremely difficult (Hettige 1990; Klein 2008). In deep soil region, soil related agricultural policies affect forests so that the clay soils are preserved for open agriculture, while slightly less fertile soils are used for agroforestry (Hettige 1990). The non-forest/forest pattern creates a patchy structure in the deep soil area, where forest is remaining in only small patches and an opposite structure in the coral rag, where the villages and shifting cultivation fields create small patches within otherwise continuous forest (Mertens & Lambin 1997). The large scale differences caused by the soil are well known and therefore not truly intriguing, but there are few large variations within the landscape regions that cannot be explained by large scale soil differences. For example why is the northern tip of Unguja dominated by semi-open vegetation instead of closed coral rag forests and scrubs? Why does the Masingini Forest Reserve still exist in the other wise deforested deep soil region? Why are there circular semi-open woodlands existing in the otherwise closed Jozani Forest Reserve? It would be interesting to study if it is the soil conditions, climate and groundwater levelz, human activity or something else that explain these variations.

There are also significant structural variations within the Unguja forests. These were highlighted by dividing forests by their structure to fully closed forests/scrubs (78% of forests) and semi-dense woodlands (22%) and based on soils and protection status to communal indigenous forests (60%), agroforests (26%), government protected forests (13%) and government silviculture forests (1%). The amount of agroforest is presumably even higher as many of these are semi-open in their structure and thus easily classified to other than forest categories. The division based on soil and protection status also partly resembles the division between forest vegetation types, as majority of those high rise mature forests that have been able to avoid severe human influence are already government protected, while the communal forests are mainly consisting of low-lying coral forests, scrubs and thickets or agroforests.

Based on the field observations about half of the closed forests/scrubs were over 5 meters in height and majority of these were within government protected areas. Although the sampling of the field observations is partly biased it seems that very large portion of closed forests/scrubs in the communal lands are less than 5 meters in height, which would mean that they would not even be forests in FAO (2000) standards. Though, they might grow over 5 meters in height if they would be allowed enough time without human encroachment. On the other hand it might be that the soil conditions are too poor and rainfall too limited for them to ever achieve this standard (Burgess & Clarke 2000: 90–91; Käyhkö et al. 2011). Even though the cover of forests is rather extensive on the island only a small fraction of these are ecologically highly valuable mature indigenous forests. It could be said that the forests of Unguja are relatively high in quantity but low in quality. Therefore measuring only the quantity of forest cover from remote sensing data is limited approach and should be fulfilled with methods able to say something about the quality.

The quality of the indigenous communal forests might be rather low in terms of forest height, but this does not automatically mean that their value is low. If values are considered from ecological perspective the indigenous low-lying forests, scrubs and thickets are part of the EACF ecosystems, they sustain diversity of tree species and provide habitats for animals (Burgess & Clarke 2000: 84–94; Kotiluoto et al. 2008; Siex 2011). In the research of Kotiluoto et al. (2008) the *shrub, thicket* and *semi-open shrubland* habitats had generally higher average amount of species and occasionally even more individual trees than the *mature forests.* Siex (2011) showed that in general there were as much traces of mammal species in the habitat type *low coral forests* as there were in *high coral forests*, while the *mixed thickets* and *shrub lands* had significantly less traces. Though, it is hard to exactly know what Kotiluoto (2008) and Siex (2011) semantically meant by their habitat types or have their sampling been spatially biased, it seems that at least the closed *low coral forest* and *shrubs* are ecologically valuable in Unguja.

For humans the low forests and scrublands are sources of many important materials and they are widely used by the local communities (Kotiluoto 2008; Fagerholm & Käyhkö 2009; Fagergolm et al. 2012). However, although they provide energy, construction materials and additional food items they are not essential for food production and are not strongly associated with non-material values (Fagerholm & Käyhkö 2009; Fagergolm et al. 2012). It might be that in the minds of local inhabitants and communities these low coral forests and scrublands are marginalized as certain kind of hinterlands, not having enough subsistence value to be considered worth developing or neither enough ecological or intrinsic value to be conserved. This might create a cycle where the scrubs are overused, because they are not considered valuable enough, which eventually prevents them from becoming valuable.

The reality nevertheless is that majority of the indigenous forests are within these marginalized forest types and their use should be organized in such a manner that it would allow increasing their quality and ecological values simultaneously as the subsistence needs of the population are guaranteed. This could be approached with better spatial planning of the land uses, especially the slash-and-burn cultivation and woodfuel collection. CoFMA process is leading the way to this as community members divide forests to “high protection” and “low impact” zones, but doing permanent plans only once is not enough as actions like shifting cultivation and wood fuel collection needs to be spatial directed almost annually (WWF Tanzania Country Office 2012: 85).

7.1.2. Changes in forest cover

Forested landscapes of Unguja have been extremely dynamic spatially and quantitatively between 1996 and 2009. About half of the forests stayed unchanged, one-third changed spatial location and one-tenth disappeared completely. The annual forest cover change rates of Unguja (-1,18% – -0,82%) are quite close to the earlier estimations done by DFNR (-1,2%) and the change measured by Käyhkö et al. (2011) from Matemwe between 1978 and 2004 (-1,14). The calculated rates are also similar to the rates of entire Tanzania (-1,16%), but clearly higher than in other African nations at the same latitude, in the other East African island nations or in the Caribbean and Oceania islands, except for Uganda and Comoros (FAO 2010b). Therefore it could be said that the deforestation situation is extremely severe in Unguja, although it does not differ from the overall situation in Tanzania.

It has been suggested that deforestation happens in a stepwise manner changing forests gradually to less vegetated and finally to permanent cultivation surfaces (Lambin 1997; WWF Tanzania Country Office 2012). These outcomes suggest similar pathways to Unguja as forests change mainly to vegetated land covers, while these change mostly to barren surfaces. However large proportion of changes from vegetated to barren surfaces have happened in well-established agricultural areas, thus are not related to deforestation. Nevertheless, deforestation seems to proceed through stepwise changed. When high forests are logged they are often also burned for cultivation (Muyuni), intensive woodfuel collection and shifting cultivation steadily turns closed scrublands to partly barren (Matemwe) or lands previously used for slash-and-burn cultivation are taken to more permanent rotation farming as population grows (Paje and Jambiani). However the current changes in Unguja rarely follow fully the stepwise path suggested by Lambin (1997) as the larger individual trees are already logged and soils rarely allow permanent cultivation.

Altogether the spatial changes happening are eventually affecting the forests more than the absolute deforestation. Within only 13 years one-third of the forests have swapped their location or in other words the forests lost somewhere have been compensated by forest gains elsewhere. The magnitude of the process is so great that it eventually influences all of the forests in Unguja. This high proportion of swapping from the total changes is rather typical for shifting cultivation and secondary forest surroundings, where areas are cleared for agriculture and let to reforest after the cultivation period (Mertens & Lambin 2000; Heinimann et al. 2007; Käyhkö et al. 2011). In their research from Matemwe Käyhkö et al. (2011) pointed out that 90% of forest changes were caused by swapping between 1953 and 1978, 74% between 1978 and

1989 and 74% between 1989 and 2004. The proportion of swapping from total changes for entire Unguja (74%) was in line with these outcomes and it seems like the process is ongoing not only in Matemwe, but all around Unguja. Certain technical aspects such as resampling, georectification, filtering and used training areas are increasing swapping, but if it is even half from the given figure the process is extremely extensive.

Eventually it is not only the easily detectable horizontal deforestation or the swapping that are deteriorating Ungujas forests, but also the continuous degradation happening vertically. Subtle degradation has been proved extremely difficult to detect with medium resolution satellite imaginary and lot of it is left undetected also in this study (Olson 1995; Lambin et al. 2003; Healey et al. 2007: 68). Majority of forests are already low-lying forests and scrubs and they are continuously degrading. In some areas, such as Kiwengwa-Matemwe and Matemwe-Nungwi the vertical degradation has been so severe that forest areas have changed their land cover class, but probably there are countless areas where forests have degraded, but not so much that it would cause land cover changes. Deforestation measured with remote sensing techniques is highly relational to the actions happening, as certain procedures such as agricultural and infrastructure expansions cause clear land cover changes, but such actions as woodfuel collection or cutting of individual trees leads to degradation that may be left unnoticed (Souza and Barreto 2000; Healey et al. 2007: 68). Therefore there is a need for systems monitoring the vertical degradation of forests, besides these remote sensing applications measuring mainly the horizontal deforestation and swapping.

7.1.3. Regional differences and spatial patterns of deforestation

As has been theoretically assumed and empirically proved deforestation is a diffusive process heavily influenced by spatial autocorrelation (Kaimowitz & Angelsen 1998: 41; Geoghagen et al. 2001; Serneels & Lambin 2001). This is the case also in Unguja as the deforested and stable forest pixels had strong spatial autocorrelation at least until 1,5 kilometers away. However the island is diverse on itself and forest changes have not been similar at different regions and parts of Unguja. It could be said that in general the government protected areas, agroforests and interior parts of the island have been more able to sustain their forest cover and large areas near and within the government protected and silviculture areas have even reforested during the study period. These outcomes support the facts that areas far from forest edges are generally safer from deforestation, government protection really functions in Tanzania and forests are less easily deforested when they create income (Ludeke et al. 1990; Mertens & Lambin 1997; Angelsen 2007: 2–4; Tabor et al. 2010). What is surprising though is that the forest cover has actually increased in the shifting cultivation area between Jozani and

Kiwengwa-Pongwe Forest Reserves. There have been plans to establish wildlife corridor in this area and the plans seems to be executing without any government support (Siex 2011).

Majority of net deforestation has concentrated into the communal indigenous forests, but surprisingly also the agroforest area is decreasing. The indigenous forests are the base of forest materials and because of all previously mentioned global and local causes it is expected to decline, but agroforests are part of income generating agricultural system and therefore not expected to deforest (Angelsen 2007: 2–4; DCCFF 2008). Urban sprawl is spreading to agroforest areas, but this process explains only a small part of the decline happening. Agroforest decline may be related to intensification of the farming practices and preferences to annual crops or methodological issues discussed later on (see chapter 7.2.4.). There have been hopes that the decreasing indigenous forest cover would be compensated by increase in the agroforest areas, at least locally (Käyhkö et al. 2011). These outcomes nevertheless suggest that this kind of process is not ongoing at the scale of the entire island.

Shifting and slash-and-burn cultivation takes place in the indigenous forests and therefore they are prone to swapping, but no surprisingly the spatial changes are more commont to agroforests, which was not assumed. There are three possible fields of explanation for these findings. Firstly, it is normal that old fruit and plantation trees are cut and new ones planted, which causes swapping, but this should not cause more spatial changes than shifting cultivation. High magnitude of these changes could be only explained by large scale government or community actions that have made the agroforests sparser elsewhere while making them denser at other locations. The second field of explanations relates to the spectral properties of semi-open agroforests. Their ground layer may vary between active and barren fields and these provide different reflectance. Semi-open agroforest with active cultivation underneath is more easily classified as forest then if there would not be cultivation happening. The swapping changes would not be then caused by changes in canopy layer, but ground layer instead. Also parts of the net deforestation of agroforest could be explained by the decreased ground layer cultivation in agroforests. The third field of explanations relate to the technicalities of classification and change detection, which are discussed later on.

More detailed spatial pattern of deforestation in Unguja can be explained rather well with the used environmental factors. This was assumed already before the study was even started as dozens of similar studies have been made around the world. Although some context related peculiarities and difficulties were found. These specialties may help us to understand the deforestation process more well at least in the highly populated island surroundings. The multivariate regression models are able to explain 75–85% of the spatial distribution of deforested and stable forests. Factors *distance to* coast, *kernel density of buildings, government protection status* and *government forestry status* where the ones that explained the process best, while *soil, mean elevation, mean NDVI, distance to roads and distance to Zanzibar Town* carried no crucial explanatory value. Although it is statistically possible to test how well variables explain the process it is completely another think to interpret why they do so.

Some studies have come to the conclusion that deforestation is more directed by accessibility than biophysical aspects (Mertens & Lambin 2000), while others argue that it is the combination of both aspects that matter (Ludeke et al. 1990; Chomitz & Gray 1996; Geoghegan et al. 2001). These outcomes however suggest that the biophysical factors have absolutely no explanatory power in Unguja. Their meaning may have been limited in other studies, but nonetheless existing. The weak explanatory power of soil and elevation is probably caused by the long human influence on the island, while in the case of NDVI the question is more about the usability of the variable.

As mentioned soil is the main factor influencing the structure of current forest cover in Unguja, but its influence on deforestation process is extremely limited (Hettige 1990). The current forest structure created by the soil differences is already so well established that large scale changes do not take place. Serneels and Lambin (2001) suggest that soils do not cause deforestation if the differences are minor and this seems to be the case in Unguja. There are no soil types within the coral rag forests that would suddenly attract farmers to clear these locations for cultivation as there are no soils in deep soil region that would be suddenly abandoned and left for reforestation, but certain fine scale differences may explain why for example shifting cultivators choose certain sites instead of others. However the used data was too robust to detect these subtleties.

Also elevation or its changes carried little or no value explaining the current deforestation process. Elevation influences deforestation process mainly when differences are significant (Kok & Veldkamp 2001). This is not the case in Unguja as the island is extremely flat. Elevation does not really make any difference, although the variable might correlate with the soil structure and explain the current forest pattern as the deep soils are generally higher in elevation. Slope was not used in this study, but I assume that also it would be connected to current forest pattern as rugged areas are left forested, but unable to explain happening deforestation (Nagendra et al. 2003)

The explanatory power of vegetation on the other hand is intriguing. Because forest vegetation type data does not exist NDVI was used in this study as its substitute. It showed ability to model the location of deforestation and stable forest in single variable regression models, but this ability was lost in the multivariate analysis. Certain forest types are more used in collection of forest materials, others provide more fertile soils for cultivation while mature forest are more easily protected (Geoghagen et al. 2001; Orjala 2008). Therefore it could be assumed that forest vegetation type influences the deforestation process, but NDVI was not able to detect the multidimensional aspects of it. Firstly, the relationship is not linear or at least not in the manner that NDVI would catch the variations. The stressed or semi-open vegetation related to low NDVI is probably vulnerable to deforestation as it has already been used for shifting cultivation, but simultaneously the lush vegetation with high NDVI value might be easily deforested because it provides better inputs for slash-and-burn cultivation and logging. Secondly, it is occasionally hard to interpret the NDVI outcomes. In Unguja the lush thickets or bushes have highest value instead of the high forest, therefore the NDVI value does not develop linearly from low stressed vegetation to high mature forests. Also the mean NDVI of the aggregated cells includes some elements of spatial autocorrelation as those cells that have already been partly cleared have lower values. To understand the inner variations of different NDVI values the variable would have required similar analysis as done for the distances. I am certain that forest vegetation type differences could help to model deforestation in Unguja, however it should be handled categorically and fed into the model separately as has been done in other studies (Boonyanuphap 2005).

In Unguja it is mainly the human influence related variables that relate closely to deforestation. The regression and distance analyses all point out that vicinity of coast increases the risk of deforestation especially in the forested coral rag region. Case studies from Matemwe and Kiwengwa show similar outcomes from a period of last 15 to 20 years (Mustelin et al. 2010; Käyhkö et al. 2011). This study shows that the process is not only ongoing in these villages, but all around the coral rag coast. There are multiple actions and multiple actors causing deforestation at coastal settings. Based on the earlier case studies and visual estimation of aerial photographs it seems that spreading of shifting cultivation is the main direct cause of deforestation, but in certain areas like in between Jambiani and Paje, Pongwe and Matemwe and in the coast of Makunduchi the spread of infrastructure such as tourist hotels, roads and houses have caused a lot of forest loses (Käyhkö et al. 2008: 76–82, 2011; Mustelin et al. 2010). Spread of tourism infrastructure also pushes local inhabitants more inland from their old coastal settlements and tempts in-migration in to the coastal villages causing deforestation in both locations (Mustelin et al. 2010; Käyhkö et al. 2011). However majority of people live by the sea in the coral rag region and therefore it is presumable that deforestation would happen close to the coastline even without the influence of tourism. Also wood extraction is causing coastal deforestation. In north of Unguja the coastal forests have been the only actual forests remaining and therefore they have faced serious pressure of forest material collection. Significant portion of coastal forest losses are also coming from the high forests of Muyuni and Uzi. Although these deforestation patches are turned to shifting cultivation fields, one could assume that monetary or material gains from the logging are also driving deforestation.

It has been generally recognized that deforestation is often taking place at the edges of the forests areas (Ludeke et al. 1990; Mertens & Lambin 1997; Nagendra et al. 2003). However the role of the coastal edge of forests in island surroundings has been less discussed matter. It would be tempting to quantify the deforestation differences of the inland and coastal edges of forests to see if there have been significant differences, but even without quantification it can be said with almost certainty that deforestation takes place mostly at the coastal edge in Unguja. However has this process been accelerated by spread of tourism, is another matter. Tourism spreads infrastructure, tempts in-migration and pushes local farmers more inland, but simultaneously it lowers the dependency on traditional livelihoods, thus possibly reducing deforestation (Orjala 2008; Mustelin et al. 2010; Käyhkö et al. 2011).

The closeness of buildings increases deforestation in entire Unguja and had especially strong role in the deep soil agroforests. It seems that deforestation caused by agricultural expansion and collection of forest materials happens relatively close to people's homes as pointed out in many researches (Mertens & Lambin 1997; Geist & Lambin 2001: 69–71; Geoghegan et al. 2001; Fagerholm & Käyhkö 2009; Fagerholm 2012). What was surprising though is that the vicinity of population had more minor impact at the coral rag region. Although the settlements are growing and causing deforestation in the region, a lot of deforestation has happened relatively far from any previous settlements. This is mainly caused by the random patterns of shifting cultivation, but might be also engaged to the spread of tourism, which seeks pristine beach locations distant from disturbing local villages (Geist & Lambin 2001: 69–71). Also some deforested areas have located quite far from settlements, but they have had to be used for woodfuel collection and shifting cultivation. The more influential role of kernel density of buildings in deep soil region could be explained with the urban growth of Zanzibar Town that has happened at the expense of nearby agroforests. All together urban sprawl has however caused less than one-tenth of all happened deforestation and the relationship is not direct or linear. As was theoretically assumed by Von Thünen already in the 19th century and empirically proved at least by Mertens & Lambin (1997), urban expansion causes limited deforestation as such, but it happens at the expense of nearby agricultural areas, which then spread outwards to the forest areas (Angelsen 2007: 2–4).

The distance to Zanzibar Town carried no explanatory power in the multivariate regression analysis and therefore it could be assumed that market accessibility as such is a limited factor explaining deforestation process in Unguja. The factor has been successfully used elsewhere, but it does not explain deforestation in Unguja as the distances in are generally too short, human presence too extensive and the soil differences too dominant (Mertens et al. 2000; Geoghegan 2001; Verburg et al. 2004). However also it may explain the current structure of land covers as the more profitable land uses concentrate to the deep soil region close to the capital as an study book example of Von Thünen's theory (Angelsen 2007: 2–4)

Another surprising locational issue was that the vicinity of roads was rather limited factor explaining deforestation in distance analysis and had absolutely no explanatory power in the regression analyses. However it has been linked with deforestation or current forest pattern in almost all modeling experiments (Ludeke et al. 1990; Chomitz & Gray 1996; Mertens & Lambin 2000; Geoghegan et al. 2001; Kok & Veldkamp 2001; Nagendra et al. 2003; Verburg et al. 2004; Boonyanuphap 2005). This can be explained by the long and extensive presence of humans, limited forest cover and diversity of land covers. Almost all the regression modeling experiments are from areas where pristine forests cover majority of the landscape and human presence is restricted to only certain locations (Ludeke et al. 1990; Chomitz & Gray 1996; Mertens & Lambin 2000; Nagendra et al. 2003; Verburg et al. 2004). In Unguja majority of current roads have been there for a long time and there is no colonialist in-migration causing serious deforestation happening along the roadsides. Shifting cultivation is dominated by the vegetation and rotation makes it to happen in a random pattern relatively far away from main roads (Geist & Lambin 2001: 69–72). Also in some areas the only remaining forests are far from roads, but the woodfuel needs have to be fulfilled anyway. However the vicinity of roads is not altogether irrelevant. The deforestation process has probably advanced along the roads in history, but this process has happened already decades or centuries ago. Also some roads and especially those that go through otherwise intact forest patches seem to cause deforestation along their sides, but the impact is so minor that it is not detected in the regression analysis. Therefore it seems like the road factor that has been successfully used explaining and predicting the spatial patterns of deforestation in highly forested surroundings does not work in areas where the forest cover is limited and population pressure extensive.

Based on these analyses it seems that it is almost solely the human related variables that explain deforestation process in Unguja, while the biophysical settings do not influence the process. The biophysical factors have created the current pattern of forests, but it is so well established because of centuries of human presence that changes are not anymore happening for these reasons. For the same reasons the road and market accessibilities, well used elsewhere, do not explain the variations in Unguja. This is only one case study from one particular context, but it might be that these generalizations could be taken to other similar surroundings with high population densities, long and extensive history of human influence, little biophysical obstacles and limited land resources.

7.1.4. Actors and actions causing deforestation

Although this research did not produce direct empirical evidence about the main actors or proximate causes of deforestation, certain links can be reasoned. Small holder farmers and urban consumers are often named as the actors of deforestation in Africa. The first ones demand land for cultivation and woodfuel for cooking, while the latter ones want food and charcoal to be provided for them (Rudel et al. 2009; DeFries et al. 2010; Fisher 2011). This is the case also in Unguja, but there are also other actors and actions in the field. Government actors have been behind forests clearances because of infrastructure development (Makunduchi), logging (Chaani) and urban sprawl (Zanzibar Town). Areas have been cleared for tourism (Nungwi, Michamwi, Paje, Jambiani and Kizimkazi) and tourism has also increased the population in nearby villages. So large areas have been cleared for shifting cultivation (Uzi, Jozani, Muyuni and Michamwi) that these actions cannot be said to be driven by individual small scale farmers, but rather by the rural communities as whole. The actors and actions behind deforestation are often diverse and therefore they should not be over simplified. It should be remembered that eventually it is not only these local actors and actions causing the changes, but there are also countless underlying causes directing their possibilities (Turner et. al 2007).

It would be interesting to study these actors and proximate causes of change by connecting expert and local knowledge to mapped deforestation clusters. These stakeholders could name all the actors and actors behind each deforestation patch and rank them by order of influence. Using maps in the process would make the experts to think the situations of each patch individually, which might provide more in-depth answers instead of simplifications. Simplifications may work at regional scale, but are rarely right answers explaining the local happenings. The actors and actions could be quantified in relation to deforested land areas, which might eventually help to rank their influences correctly.

7.1.5. The underlying causes and future of the forests

There are multiple ways to calculate the annual deforestation and forest cover change rates (FAO 1995; Puyravaud 2003). Though the interest in here is not in the calculation methods as such, but rather in the future scenarios different rates produce for the forest cover of Unguja. The figures drawn from the gradual classification suggest milder change (5,4 km^2 & -0,82%), while the figures of the abrupt classification are more severe (7,4 km^2 & -1,19%). The generally used annual forest cover change calculation methods are constructed so that they assume that the deforestation declines along the declining forest cover (FAO 1995; Puyravaud 2003). Therefore these percentage figures -0.82% and -1,19% would level the absolute area lost annually along the time, while the average annual absolute deforestation figures (5,4 & 7,4 km^2) argue that change keeps its current pace. When projected until 2025 the annual forest cover change rate calculated from gradual classification suggest the highest forest cover (501 km^2), the average annual absolute deforestation figure from the same classification and the annual forest cover change rate drawn from the abrupt classification come to mediocre outcomes (472 km^2 & 485 km^2), while the average annual absolute deforestation figure of the abrupt classification suggest the lowest forest cover (454 km^2).

There are few methodological issues supporting the scenarios build from the modest figures of the gradual classification. Firstly the abrupt classification exaggerates deforestation, as it classifies scrublands as forest in the SPOT image. Although, the gradual classification on the contrary underestimates the deforestation as changes from forested classes to *semi-open scrubs on grass* are left outside, the problem is more severe to the other direction. Secondly the data used for calculating the change rates was not filtered and significant proportion of deforestation is caused in really small forest patches, which would not be even considered as forests because of their size in FAO (2000) standards.

Besides the methodological issues there are few matters related to the local context, which promote the use of more moderate deforestation estimates. Firstly, a significant portion of change has been caused by government actions. The silviculture area of Chaani has been cut and large areas from Makunduchi have been cleared for urbanization. It may be that similar actions are taken also in the future, but all of the government forestry sites are in the stage of regeneration at the moment and it is likely that they will rather increase the forest cover instead of decreasing it. Also the willingness of people to resettle in planned areas like Makunduchi seems limited in the light of buildings built there and therefore it seems unlikely that similar actions are promoted in the future. Secondly, in some areas, such as Fumba peninsula, Uzi Island and near Nungwi and Matemwe, the forests have disappeared almost completely, therefore removing the base for deforestation. The subsistence need for materials like fuel wood has not disappeared and will be partly fulfilled by cuttings in other parts of the islands. However the nearest forests are far from these locations and therefore I expect that the need for wood fuel is fulfilled with overuse of existing semi-open scrublands, as has happened in Matemwe (Käyhkö et. al 2011). Thirdly, although the tourism expansion is probably going to continue in the coastal areas, the boom of it has happened during this time period, therefore the tourism related coastal deforestation should be slowing down (Gössling 2001; Mustelin et al. 2010). Fourthly, the trend of deforestation is declining at least in Matemwe (Käyhkö et al. 2011). Though this is only a single case study and its outcomes are caused by local context and therefore should not be generalized for the whole island.

Eventually it is the underlying causes that specify deforestation happening in the future It is important to estimate has the time period used to determine the change rates been somehow exceptional and how changes in underlying causes are going to influence the future changes (Kaimowitz & Angelsen 1998: 90–98; Geist & Lambin 2001: 6–8). At global level deforestation trend seems to be slowing down, but accelerating in Africa (FAO 2010). Some have even argued that deforestation has not really even started in Africa and logging, export agriculture and bioenergy driven deforestation will skyrocket the deforestation figures in the following decades (Rudel et al. 2009; DeFries et al. 2010; Fisher 2010). However I consider Zanzibar to be largely safe from these actions as its land and forest base is so small and island location makes transportation costs outrun the possible incomes. Therefore it is likely that deforestation continues to be subsistence agriculture driven as it has been so far in the island and generally in Africa (Käyhkö et al. 2009; Rudel 2009; Siex 2011). Although it is possible that incomes provided by cultivation of bioenergy species that are able to grow in poor soil conditions, such as *Jatropha curcas,* would promote clearance of the coral rag scrublands.

Even though the main underlying causes of deforestation would not be changing the amount of persons needing subsistence is going to increase. The population more than doubled between 1978 and 2002 and it will probably do so between 2002 and 2025. The annual population growth rate of Unguja was 3,2% between 1978 and 1988 and rose to 3,5% between 1988 and 2002 (OCGS 2007: 10). Even if the annual population growth rate would not be increasing the absolute population is going to do so for a long period. Large proportion of this population is going to get their livelihoods from agriculture and pressure for permanent and slash-and-burn cultivation of forests increases. Also 96% of the households use charcoal or wood fuel as their main energy source for cooking. Even though DCCFF (2008) uses more moderate estimations for population growth, it has assumed that the domestic demand for wood increases 44% until the year 2020. Alternative energy sources, such as liquid gas and solar cookers, more efficient use of wood energy through fuelwood saving stoves, awareness raising and tree planting are planned and proposed solutions for the growing energy demands (DCCFF 2008; Käyhkö et al. 2009; Fagerholm 2012: 53). However thousands of new equipment should be taken in to use to even keep up with the growing demand.

The population growth in the West region, where parts of Zanzibar Town and its conjunctions are located, has constantly been higher and has also accelerated faster than elsewhere on the island (OCGS 2007). The rapid urbanization may relief the deforestation pressure caused by the rapid population growth and the agricultural and wood energy demands related to it. As mentioned generally urban dwellers are straightaway off from agricultural activities, they are more prone to take new energy solutions into use, their demand promotes tree planting in rural areas and they use relatively more food and energy products imported from elsewhere. (Contreras-Hermosilla 2000: 21; Foster & Rosenzweig 2003; Rudel et al. 2005; Writght and Muller-Landau 2006; DCCFF 2008). However Zanzibar Town also tempts immigrants outside the island, urban dwellers use more agricultural and forest materials than their rural counterparts, inefficient charcoal is used in cooking, they have limited means to participate in wood energy and agricultural production and the spread of the city causes direct deforestation (Masoud 1991; OCGS 2007; DCCFF 2008). Also the forest products they use are not coming from their backyard, which weakens the human-environment relationships and makes it easier to make environmentally unsustainable consumption demands. Eventually it is difficult to say has or will urbanization reduce deforestation in Unguja as it depends how consuming habits change, where the food and wood materials are produced and does urbanization tempted migrants outside Unguja. However I assume that the urbanization has and will help to mitigate the deforestation impacts of the population growth.

Besides these issues it should be also acknowledged that the study period has been economically and politically abnormal. The economy has been liberalized and annual GPD growth has been impressive (RGZ 2009). As mentioned in the theoretical section economic growth has many positive and negative effects on deforestation and in Ungujas context the growth has been driven by the tourism sector (RGZ 2009; Gössling 2001). The economic growth may have promoted forest clearance to make room for infrastructure developments, urban sprawl and tourism, but also provided resources for management of the government forest areas and creation of new protected areas. Zanzibar also faced political turmoil causing foreign governments to stop their development projects in 1990s. This period also created a caesura in the well-established collaboration between Finnish and Zanzibar governments in forest sector, decreased foreign funding of DFNR and impaired the possibilities for sufficient forest management. However partly because of the previous collaboration the country had governmental forest management actor existing during this study period (Sitari 2005).

It might be that the current common land tenure system, where land as such is not owned, but the assets on it are promotes actions that cause deforestation (Törhönen 1998; Fagerholm 2012: 34–35). Forests may be overused in shifting cultivation and forest material collection as people try to maximize their personal benefits (Hardin 1968). Also certain kind of land speculation might exist. Forests in potential tourist areas might be turned to fields or habituated to get income from these lands if they are later sold for tourism use. Also high mature forests might be turned to shifting cultivation fields if there is a risk that government seeks to turn these to official protection areas (Conteras-Hermosilla 2000: 16). These are solely assumptions, but the current land tenure system supports these kinds of actions and any changes in it would inevitably affect the deforestation process.

All these factors have had essential influence on forest changes happened during the study period, but their influence has been so diverse and interlinked that it is impossible to estimate how much, why and to what direction the effects have really been (Geist & Lambin 2001: 1–2; Lambin et al. 2001). Also it is extremely difficult to estimate how these issues are going to influence forest changes in the future. However after acknowledging all these underlying causes and assuming that radical changes are not going to happen in near future it is likely that the average annual area deforested measured from the gradual classification (5,4 km^2) creates the most accurate future estimations, at least until 2025. Because of the rapid population growth, there are no reasons to believe that the absolute area lost would be decreasing, but it should be acknowledged that the abrupt classification gives exaggerated figures. It is also possible that population growth, in-migration, economic development and spread of tourism in the forest rich coral rag region might actually even increase the annual deforestation in the future.

However where this future deforestation takes place is completely another question. At the moment deforestation is highest near build-up coastlines and if the situation continues as such the still remaining coastal forest are at high risk of deforestation. From all of the endangered coastal forest, there are two patches that stand out as especially valuable and vulnerable: one in the southern tip of Uzi Island and one in the west coast of Muyuni (Figure 33). Two reasons make these patches rather unique: Firstly, they are the only high forest patches without government protection at the direct vicinity of the coast and secondly, they have already faced serious deforestation. Aerial photographs and GeoEye images reveal that the deforestation has happened between 2004 and 2009 and if this pace keeps continuing there is nothing left of these patches within next five to ten years. The monetary and human resources of forest protection are scarce in Zanzibar, but from my opinion the ecological, environmental, aesthetic and cultural values of these patches should be studied and they should be taken under official government protection before it is too late. The patch of Muyuni is in somewhat better position than the one in Uzi, since it has been included to the original plans for future network of government protected forests as a part of “Muyuni-Jambiani” forest conservation area, but in the last delineations of the new protected area provided by DFNR, this patch is missing (Siex 2011; WWF Tanzania Country Office 2012). Both of these patches are also marked as “High protection zones” in CoFMA agreements, but the authority of these plans can be questioned, since parts of the Muyuni -patch delineated under protection has already deforested since 2004.

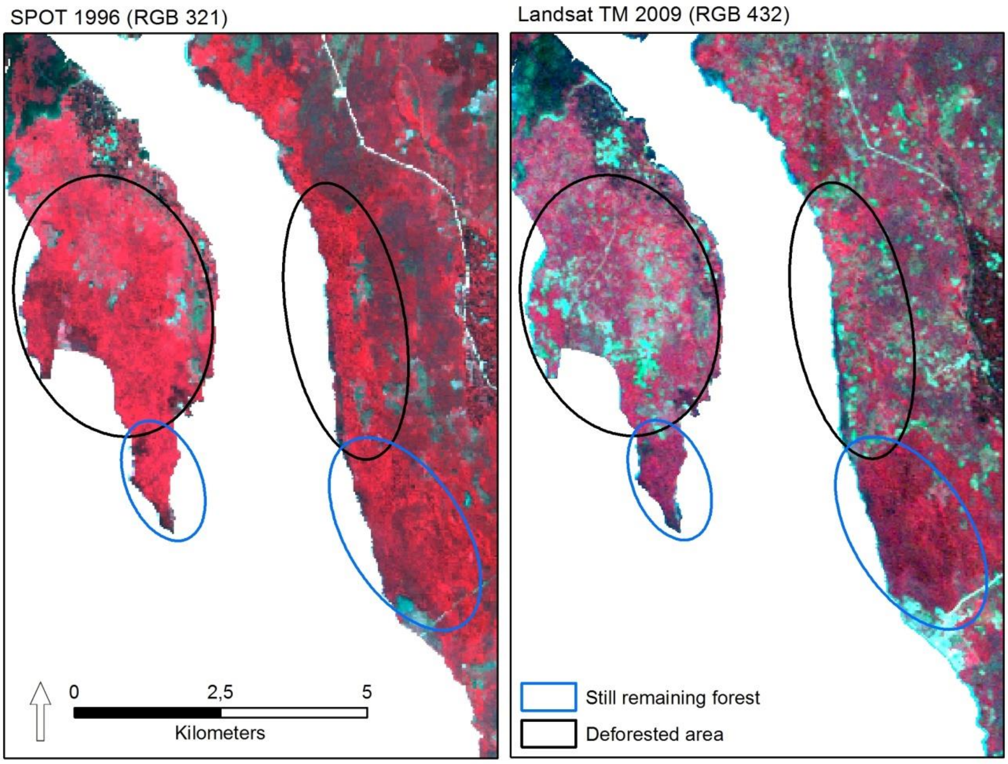


Figure 33. The forest areas lost between 1996 ans 2009 and still remaining forests areas under deforestation risk in Uzi and Muyuni.

Another area that is under severe risk of future deforestation is the agroforests close to Zanzibar Town (Figure 34). Deforestation caused by urban sprawl is a complex issues as there are tens of thousands individual actors and also government and city administration involved, but if deforestation keeps its current pace are the nearby agroforests seriously threatened.

A third set of areas where the development of deforestation should be more closely monitored are the outskirts of coastal villages in the South district. Slash-and-burn has been the main cultivation method there and more areas are cleared because of this today than in 1996. However it seems like other more stable forms of rotation farming have also been established during the study period. There are clear marks of this close to villages of Bwejuu, Paje, Jambiani and Kizimkazi (Figure 35). The population has grown significantly in these villages because of natural population growth and tourism related in-migration. It may be that, because of the population growth people have had to harness livelihoods that were earlier not used in the area. However these farming methods may be unsustainable in a long run, which may lead to similar degradation of forest that has happened in Matemwe and to a cycle where agriculture deteriorates the soils and forests of a site and then spreads to new areas (Käyhkö et al. 2011).

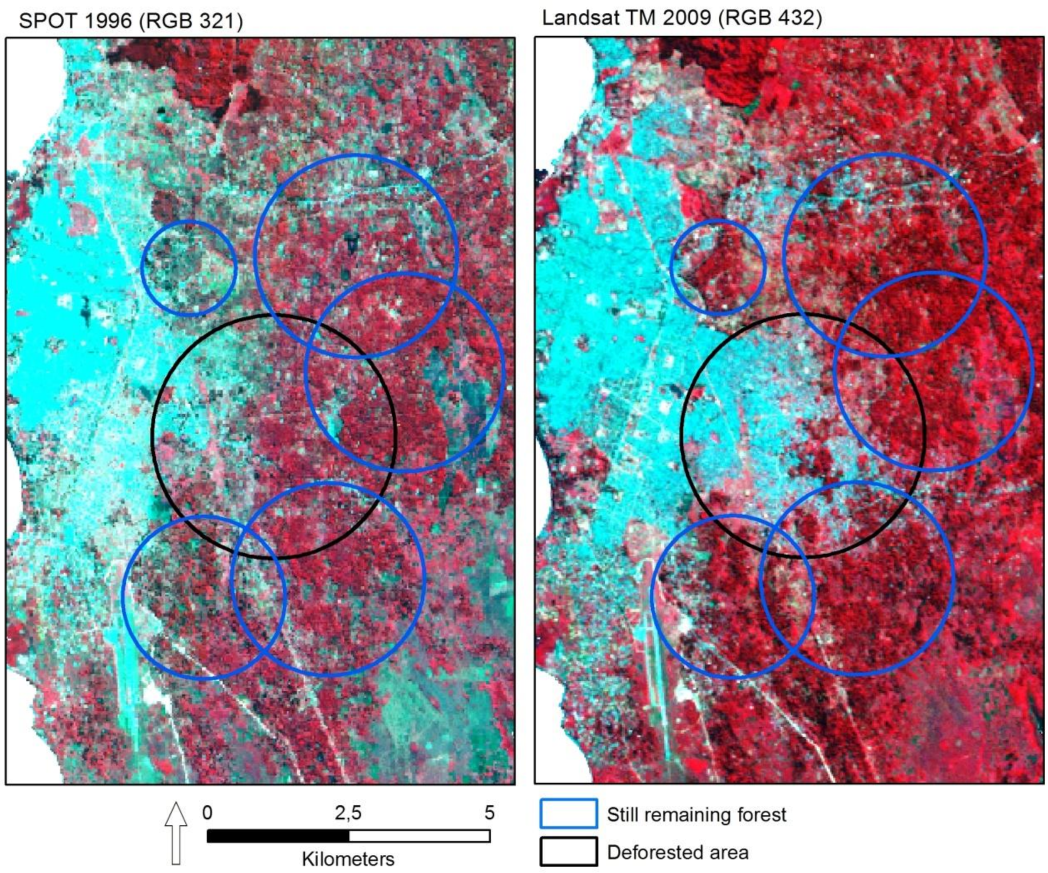


Figure 34. The agroforest deforested between 1996 and 2009 and forest patches under deforestation risk near Zanzibar Town.

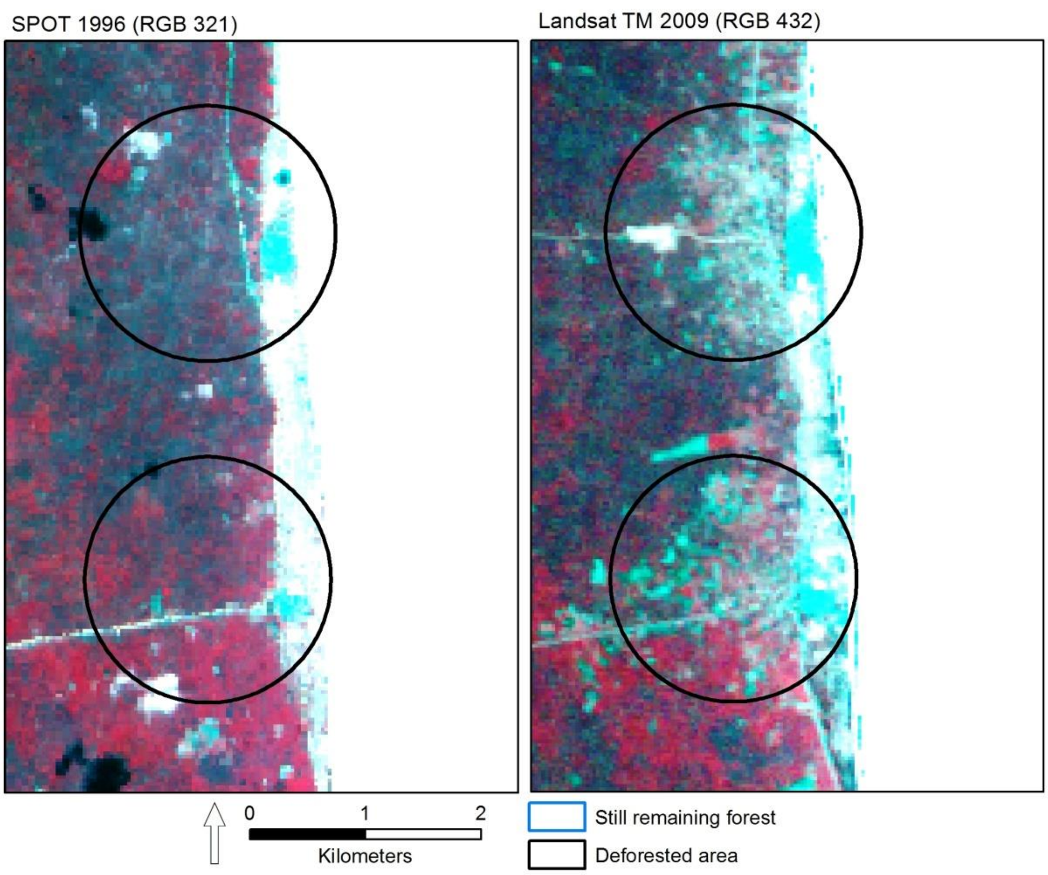


Figure 35. The spreading of farming near Bwejuu and Jambiani villages between 1996 and 2009

It is likely that in a long Zanzibar is following the “forest scarcity path” of forest transition where decrease first slows down and then turns to increase, because of the increasing forest product prices (Rudel et al. 2005; Barbier et al. 2010). However the moment when the deforestation turns to reforestation or even slows down is far from the present. It is impossible to estimate when the forest cover reaches its bottom and how much there would be forests still existing at that time, but the latter makes all the difference from the perspective of biodiversity. There is not much hope for the flora and fauna if the forest cover drops to 1% before starting to increase, but the outlook is not as bad if 15-25% of the land can be kept forested. The question is not only about the quantity, but also about the spatial pattern and quality. Percent-wise even relatively small amount of forest habitats can help to sustain the biodiversity if it is connected, diverse and in good conditions (Jongman & Pungetti 2004: 2–33). Therefore I see it extremely important to promote the already made plans to establish a network of protected areas that would create one new conservation area in to the south and connect all of the major forest areas with wildlife corridors (Siex 2011; WWF Tanzania Country Office 2012). However the current spatial plans for this network are rough and they should be optimized so that the spatial pattern and diversity of habitats would ensure maximum biodiversity benefits with minimum costs.

There are also few relatively far-fetched, but existing possibilities that could make the “economic development path” of forest transition possible or on contrary risks that could lead to rapid decline or even devastation of current forest base. Oil and natural gas has been found from the offshores of Zanzibar. As such, the use of oil and gas could change the energy sector of the island and practically abolish the large-scale need for woodfuel as has happened in some oil rich countries, but it would also bring in the other benefits of economic development. Improved economic situation could relief large amount of rural inhabitants from agricultural subsistence economy, connect them to monetary economy, increase the importation of food products and promote environment protection attitudes, spatial planning and tree planting (Rudel et al. 2005; Barbier et al. 2010). It would eventually eradicate the need for ecologically destructive and unprofitable subsistence actions as the slash-and-burn cultivation. However, the influence of the economic development is highly related to local social, political, institutional and cultural context (Conteras-Hermosilla 2000: 19; Scrieciu 2006; Barbier et al. 2010). Eventually how the economic development gained through oil and gas will change the forest cover in Zanzibar is not about the absolute economic growth as such, but rather about its equalitarian distribution among the population and how the money is used to build the structures of the society and especially those of environmental management. Following the unequal development paths of Nigeria or

Angola would not change the ways impoverished rural inhabitants use their surrounding environment, at least not to positive direction (Omotola 2006; Maconachie et al. 2009).

From the negative side it should be mentioned that the food self-sufficiency is only around 50% in Zanzibar. The growing population may lead to increasing soil erosion and degradation in agricultural lands, which would decrease the already low food self-sufficiency (RGS 2004). Also as the population grows and urbanizes more of the forest materials are brought from mainland Tanzania or elsewhere. If importation of food and wood from mainland would be disturbed, possibly because of global food crises or political tension caused by ever growing independency demands, this might cause severe destruction of the unprotected communal indigenous forests as desperate needs would be taken to ensure food and energy availability. Long-lasting crisis could lead to severe deforestation and devastate the current forest cover as has happened in North Korea (Noland et al. 2001). As already mentioned the spread of bioenergy cultivation in coral rag lands could also lead to severe decrease of forest cover.

7.2. Methodological issues

7.2.1. Errors in the preprocessing

The errors in the georectification did not really influence the outcomes of the individual classification of 2009, but caused displacement errors in the change detection. The general RMS error of was over the 0,5 pixel, which is often considered as a critical line in change detection (Lillesand et al. 2008: 595). It seemed that if the northern end of the island was well rectified there were serious errors in the south and wise-versa. In the final rectification the errors are greater in the northern end of the island. The original raw data was provided in two pieces and the north and the south parts were originally different files. There was approximately one pixel mismatch in the borders of these two original images, which disappeared when the images were mosaicked. It might be that this original mismatch prevailed in the data and made adequate rectification impossible using polynomial equations. This problem could have been avoided by first matching the images in a regular graphics software before mosaicking and georectifying them in a geospatial software. Another dilemma in the rectification was that although it is generally recommended to rectify the low resolution images against the high resolution ones, this was not possible in this study, because the Landsat TM file was already rectified against other used data (Mather 2005: 88). It was tested to rectify the SPOT image against the high resolution aerial photographs used to rectify the Landsat image, but the outcomes were even worse.

7.2.2. Challenges of land cover classification

The created classifications are governed by use purposes, available datasets and possibilities for field working. The classification process would have been different if it would have been done only for the 2009 image, but as it was done also for the 1996 image and both of the classifications were used in the change detections the classification solutions had to adapt. Using solely the supervised method for a single classification would have been preferred as it creates ready classes, while on the other hand the unsupervised technique would have been more adequate for the change detections as it relies on statistically sound clustering of the spectral values and is therefore less prone to human errors (Campbell 1996: 317–329; Lillesand et al. 2008: 557–569). It is possible to misinterpret the spectral clusters afterwards and some of them are meaningless in the sense of land covers, but at least they are always spectrally coherent. In supervised classification it is possible to create classes, which have so wide spectral range and standard deviation that they capture pixels not belonging to the particular class (Lillesand et al. 2008: 557–569). In the created classifications the spectral land cover features *semi-open scrubs on grass* and *dense herblands* and the final land cover class *semi-open scrubs on grass* faced these errors, which then had to be minimized by creating two different change classifications.

Using solely unsupervised classification was not possible in this study as there was no spatially and temporally adequate reference data for the 1996 SPOT image that could have been used to connect the spectral clusters to land cover classes. The aerial photographs used as reference data in accuracy assessment were from 1988–1989 and during this 7 to 8 years temporal gap a lot of land covers had already changed. The images were also black-and-white making interpretation of certain features such as the bareness of the ground layer or forest vegetation type extremely difficult. Also the images covered only certain areas of the island and although there was a lot of on-site knowledge about these areas based on previous research, some land cover types were not represented in them (Sitari 2005; Käyhkö et al. 2008, 2011, Mustelin et al. 2010, Fagerholm et al. 2011). Eventually also the statistical nature of the used data would had influenced the unsupervised classification so that created clusters might have not been completely comparable between the images. For example parts of the low-lying scrubs could have been joined with the semi-open scrub classes in the other classification, while they could have belonged to some forest class in the other (Campbell 1996: 317–319; Lillesand et al. 2008; 568–570).

The focus was set to minimizing the errors of the supervised classification as the unsupervised technique turned impossible. The idea was to make the training sites as homogenous as possible (Lillesand et al. 2008; 557). One way to ensure this was to delineate the training sites based on the unsupervised clusters. However this was done only for the 2009 classification. Although it would have been impossible to interpret the spectral clusters from the 1996 image, the unsupervised clustering could have been used to see that are the training sites homogenous and that has there been significant changes between 1996 and 2009. Now these issues were assessed based on the 1988-1989 aerial photographs and histogram matched SPOT image. It was possible to detect major changes with these datasets, but subtle changes were left unnoticed as certain thinks simply looked different in the SPOT or the black-and-white images. For example some forests looked significantly brighter in the SPOT image than in the Landsat scene and it was hard to decide has the vegetation within the training site been lusher in 1996 or were these visual differences simply caused by the sensor calibration. Eventually these subtle disparities within the training sites may have caused significant differences in the final signature files, thus causing wrongly detected changes. Some of these errors were minimized by using four different land feature classes for forests that all had different spectral responses and post-classifying these to one forest land cover class only after the automated classification process.

As the accuracy assessments shows the final land cover classifications work rather well for the forests and urban areas, but are inaccurate for other land cover types. The inaccuracies are largely due to the limitations of the used spatial resolution and classification method. The pixel size can be simultaneously too coarse, adequate and too detailed. It may be able to detect certain things, while simultaneously unable to detect others (Campbell 1996: 316; Di Gregorio 2005: 1; Lillesand et al. 2008: 110–116, 617–621). The landscapes in Unguja are extremely heterogeneous, small-scaled and mosaic-like (Klein & Käyhkö 2008). There can be multiple land covers even within the 30 m^2 pixels and majority of pixels are mixels of multiple spectral responses (Foody 2004). Urban areas and forests were correctly classified, because their real world spatial and spectral properties correspond with the used pixel size. In other words they are usually spectrally rather homogenous and spatially continuous within the 30 m^2 pixels. On the other hand certain land cover types such as semi-open areas are neither spectrally homogenous nor spatially continuous. Within large semi-open areas pixels can be aligned so that there is completely closed canopy coverage within one pixel and absolutely no canopy in another (Figure 35). Therefore the used pixel size is both too coarse to detect the individual components that make the land cover, but simultaneously too detailed to detect the spatial pattern of the land cover in focus. These problems could be partly solved with spatial pattern recognition or object-oriented classification techniques acknowledging not only the spectral properties of a cell, but also the spatial patterns surrounding it (Lillesand et al. 2008: 545–591). Use of these methods would take an important step away from spectral land covers towards more cognitive land covers.

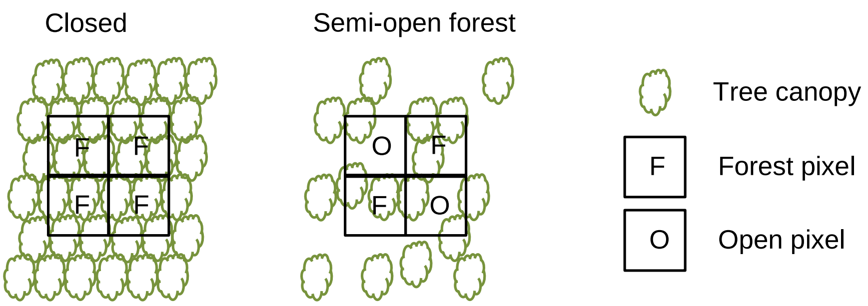


Figure 36. Conceptual model about the difficulties of spectral land cover classification in semi-open surroundings.Some of the pixels in semi-open areas get their reflectance from ground layer and other from canopy and they are therefore classified differently although they would belong to same land cover class.

Although the forests are generally classified quite correctly the class as such is a crude generalization. As mentioned it includes agroforests, thickets, scrubs, low coral forests as well as high mature forests all formed from various species with different spectral responses (Campbell 1996: 316; Di Gregorio 2005: 1). The original idea was to create classification, which could have detected few main forest structure or vegetation types. This proved more difficult that thought and the four forest land cover features in the original classification were post-classified to only two for the final one. Also it would have been rather impossible to make sure that the land cover features represent same forest types in the SPOT image as there were not adequate reference data. However there a countless studies were different forest types have been separated using Landsat TM imaginary and there are no reasons why this could not be possible also in the context of Zanzibar (Foody & Hill 1996). One approach to improve the created forest classification would be to clip the original data with the created forest borders and run the unsupervised classification only within the forest areas to see the inner variations (Coppin et al. 2004).

Eventually what is considered as a “forest” is a problem in any forest cover measurement. Creating conceptual classes may be a start for a solution, but the problems begin when the concepts are taken from abstract level to reality and boundaries between land cover types are drawn (Ahlqvist 2004, 2008). FAO (2000) have created conceptual description for forests, but this was seen poorly applicable for this study, mainly because it argues that the canopy coverage in forests should only be over 10% and trees should reach 5 m height in maturity. If the 10% coverage -rule would have been used majority of Unguja would be classified as forest, while if the 5 m height -rule would have been used large part of the indigenous coral scrubs would not be considered as forests. Also it is extremely difficult to detect forest heights from optical remote sensing data (Garcia et al. 2010). As these predefined descriptions are badly applicable for local contexts and remote sensing data, research often relies on subjective descriptions created by researchers themselves case by case (Weiers et al. 2002; Ahlqvist 2004; Olander 2008). However these create subjective outcomes and comparing forest cover estimations created by different people at different times for different locations is extremely difficult and often biased (Ahlqvist 2004, 2008).

Methods have been developed to back-up the subjective classes with qualitative and quantitative descriptions, such as general land use, visual appearance, average canopy coverage and photographs (Ahlqvist 2004). This research tried to open the semantics by attaching the collected field data to the classes, but as this data was used also to direct the classification the outcomes are partly biased. It also became obvious that there should be enormous amounts of field observations to create representative descriptions. Therefore I consider creating semantic descriptions based on the used raw data as more adequate approach for remote sensing based classifications. Mean NDVI was attached to each class partly for this reason, but knowledge about spectral class centers, variances and standard deviations in all used bands should also be somehow attached to the classifications. This information would allow other user to recreate the signature files used in classification and repeat the classification perfectly.

What comes to the accuracy assessment of the final classification outcomes there is a serious reason to believe that the done accuracy assessment underestimates the error in the data. Firstly the accuracy was only assessed from those aggregated cells that had variance value one meaning that there was only one land cover type within them. However these cells only resemble the areas where the land cover is spectrally and spatially homogenous, while the most severe accuracy errors are in the borders of classes or within mixed areas (Lunetta & Lyon 2000: 6–7; Foody 2004; Lillesand et al. 2008; 587). On the other hand it might be impossible to identify the absolute land cover of border areas even with cognitive human vision as these areas are often fuzzy and gradual. Nevertheless, as these areas were avoided in the assessment the final accuracy figures are higher than they should be. I would estimate that the figures are at least 10 to 20% too high for each class and for the entire classification.

7.2.3. Uncertainties related to change detection

If creating the classifications had severe uncertainties related to the used method, class semantics and accuracy assessment, change detection is facing even more challenges. Many of these are normal for any change detection studies, some are emphasized by cross-sensor analysis and few are typical only for the cross-sensor analysis. The largest errors in automated change detection techniques comparing classifications pixel-by-pixel is the comparability of these pixels (Coppin et al. 2004; Mather 2005: 88; Wulder et al. 2008). The georectification caused change detection errors in two different ways, by influencing the comparability of the training sites between the raw images and by influencing the comparability of the pixels in the final classifications. The slight differences in georectification and resampling made it impossible to say were the training sites in the same geographical location between the two images, which increased the differences between the two classifications (Campbell 1996: 327–329; Mather 2005: 86–90). This problem was reduced by making the training areas smaller from their edges so that possible mixels and rectification errors would be avoided. Unfortunately the landscapes in Unguja are so heterogeneous that it is difficult to create large continuous training sites and thus majority of the sites had little to reduce from.

It is hard to estimate how much wrongly detected change it caused that the pixels were not perfectly aligned because of the insufficient georectification. However it is likely that the errors were larger in agroforest than in other forest areas. Agroforests in Unguja are relatively small in size and they form semi-open patterns where the forest pixels mix with other land covers, while the indigenous and government forests are relatively large in size and continuous in their spatial pattern. If there is a situation where forest cover have not changed at all, but there is a one pixel georectification error the change outcomes are significantly different for continuous and semi-open patches. In continuous patches the change error only occurs in the borders of the class so that deforestation on the other side is compensated by reforestation at the other border, while the interior forests stay unchanged (Figure 37). However the outcome may cause all the pixels in semi-open areas to change if the pixels are aligned so that each forest pixel is neighboring an open pixel. This partly explains the extremely high swapping of agroforests.

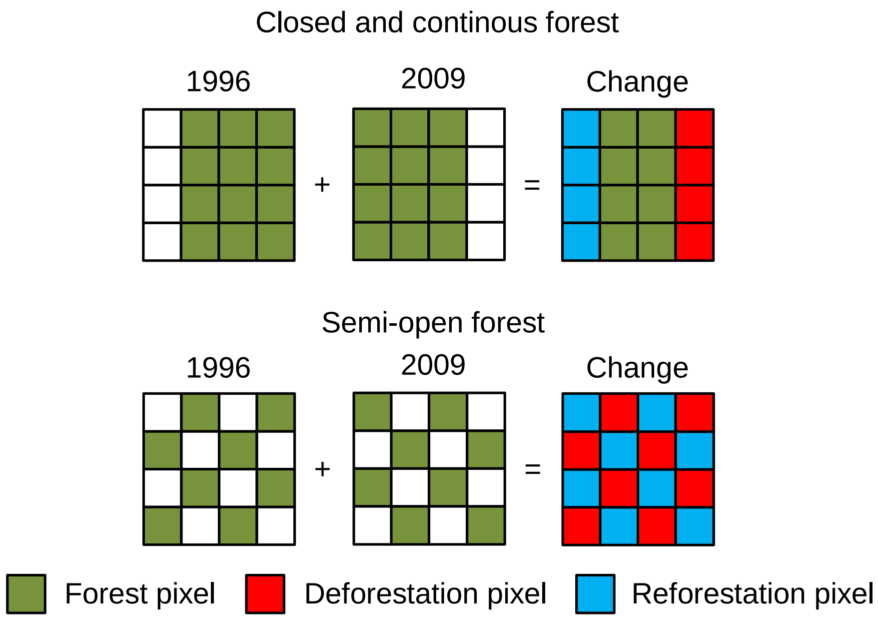


Figure 37. Conceptual model about the change detection errors caused by geo-rectification in closed and semi-open forests. Rectification causes errors only in the borders in the closed forests, but all the pixels may be affected by the error in semi-open forests.

Second technical issue influencing change detection relates mainly to cross sensor analysis with different spatial resolutions. The Landsat TM had original pixel size of 30 m^2, while this was 20 m^2 in the SPOT image. The SPOT image had to be resampled during the georectification process to match the Landsat resolution and this was done with the nearest neighborhood technique. Resampling with neighborhood technique reduces the amount of data as it assigns the pixel value from the nearest cell center of the original data to the new resampled pixel. When data is resampled to larger pixel size it means that certain pixel centers are not closest to any of the new pixel centers and therefore their pixel values are not attached to any new pixel, thus reducing the amount of information (Lillesand et al. 486–490). The reduced data is not a problem as such, but in certain cases it may influence the change detection outcomes. The conceptual model in Figure 38 shows a situation where the actual forest cover stays stable, but neighborhood resampling causes change detection errors. The numbers 1 to 4 and letters from A to D represent the SPOT pixels. The pixel in the upper left corner is pixel 1A and the one in lower right corner is 4D. The small letters a, b, c and d represent the Landsat pixel they are inside of and the arrows show which the SPOT cell centers are used to assign data values in the neighborhood resampling. As it can be seen the Landsat pixels a and d have over 50% forest cover and are therefore classified as forests, but when the SPOT image is resampled to matching spatial resolution these pixels are classified as non-forest because the forest coverage in the SPOT pixels closest to new cell centers (2B and 3C) is significantly less than 50%. The situation is opposite for pixels b and c. This causes the pixels a and d to be classified as reforested and pixels b and c as deforested, although the situation has stayed stable in reality. Also this phenomenon partly explains the high swapping and net deforestation of the agroforest areas as the semi-open forests are more prone to these kinds of errors than the continuous indigenous forests.

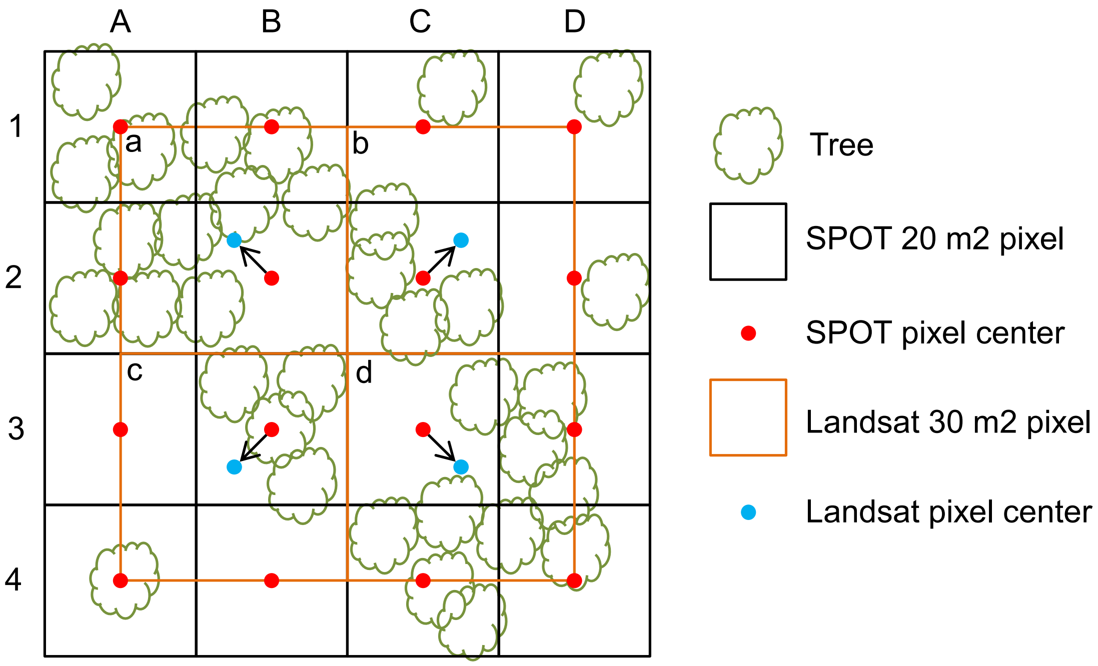


Figure 38. Conceptual model about the change detection errors caused by resampling to larger pixel size with the nearest neighbor method. The smaller SPOT pixels used have different forest cover than the larger Landsat pixels they are assigning values, which causes change even though the situation has stayed stable.

It would be scientifically tempting to empirically measure what is the general influence of the used cross-sensor resampling method in change detection. This could be done by having SPOT images with various spatial resolutions and a Landsat TM or ETM+ image from the same time and the same area. In the experiment the SPOT images would be resampled to Landsat resolution with different methods, the images would be classified with the same training areas and the change caused by resampling in theoretically same scenes would be measured with normal post-classification comparison techniques. Only problem is that the resampling is so closely connected to done georectification, that in practice it would be completely impossible to say that the pixels in the images would actually represent the same areas (Lillesand et al. 2008: 486–490).

For these and other reasons the change detection should not focus to individual pixels, but rather to continuous patches (Wulder et al. 2008). This could be achieved by filtering, aggregating, amalgamating or generalizing the original or the final change classifications (Coppin et al. 2004; Lu et al. 2004; Franke et al. 2006). However none of these procedures where done for the original change data and therefore the non-filtered data is also present in the change rates, KDE, sub-area and distance analysis, while the regression and predictive analysis were done with aggregated data. There are reasons why these methods were not used. Filtering was seen too harsh and influencing the final outcomes more than the presence of small patches as it simply changes pixel values based on their neighbors. The outcomes of filtering would have been highly dependent on the spatial pattern of original values, which might have caused serious errors in the agroforest areas. Certain amalgamating filtering on the other hand could have been used if there would have been options either in ArcGIS or in Erdas Imagine to merge individual pixels to surrounding mosaic when 6–7 of its neighbor pixels would belong to same category. However the softwares did not offer these kinds of solution and it was seen too complicated to build this kind of application by one's own. Option that should have been used would have been to delete the smallest *deforestation* and *reforestation* patches after the change detection. Deleting the patches smaller than the FAO (2000) 0,5 hectare limit for forests would have reduced the deforestation estimations approximately by one-third. However the FAO (2000) limit is a minimum size for forest patches and not for deforestation patches and properly using this limit would have not meant deleting all deforested patches smaller than 0,5 hectares, but only deleting them from forest areas that have had original size smaller than that. All these generalization procedures could have helped to remove some of the errors caused by georectification, as some of those small patches which were more prone to changes caused by rectification errors would have been removed (Lillesand et al. 2008: 595; Wulder et al. 2008).

Also a likely source of inconsistency was created when the final supervised classification was done with six spectral bands for the Landsat TM image and only with three bands for the SPOT image. In theoretical terms the Landsat TM had a wider spectral resolution, whereupon it created the classes differently (Lillesand et al. 2008; 411, 433, 554–557*)*. It may even be that certain Landsat pixels had exactly same reflectance in the three bands used in both classifications, but they were classified differently, because of the differences at those three bands used only for the Landsat TM image. The bands that were only used for the Landsat TM image were able to detect urban elements, soil and vegetation moisture differences (Lillesand et al. 2008: 411). This explains why the SPOT classifications underestimated the urban areas. Also some closed scurblands with low moisture levels could have been classified as semi-open instead of forested, which might explain why there were so much swapping between semi-open scrubs, woodlands and forests. Especially the areas analyzed as deforested in the northern parts of Unguja look relatively similar between the raw images with pseudo-colors and I assume that it is partly these moisture differences that caused the wrongly detected change.

Eventually these issues related to sensor calibrations, amount of bands used, georectification and resampling influenced the training sites collected, the signature files created and the final classifications produced. Even though the semantic concepts of the classes are equal between the classifications, their technical executions are not and this creates errors in the change detection (Campbell 1996: 327–329; Ahlqvist 2004; Lillesand et al. 2008; 411, 433, 485–490; 554–557). In this study, especially the transitions between *semi-open scrubs on grass* and the two forested classes were often considered unreliable based on visual estimations. Changes in visual appearance between the classes were so subtle that one could not interpret that were they caused by actual changes in vegetation or just by technical-semantic differences. These subtle vegetation differences could be caused by logging of larger trees, coppicing of branches, intensified shifting cultivation or annual changes in rainfall and climate. All of this could be interpret as natural degradation. Another explanation is that the pixel values in these areas set close the class borders and minor differences in the training areas cause them to be classified differently between the images. It might be also possible that the differences in the spectral resolution caused these errors. Whatever the explanation is, these problems made it obligatory to create two different change classification schemes, which eventually led to range of forest change rates instead of one precise figure. The gradual classification underestimates the happened changes, while the abrupt classification overestimates it and the real figures are probably somewhere between these estimations. Although, this helped to reduce the change detection errors, the problems were so substantial that the outcomes of this study should be interpreted rather as suggestive than as absolute truth.

However many of these dilemmas could have been avoided by using data from a single sensor, such as Landsat TM (Wulder et al. 2008). Firstly, the Landsat TM is in one spatial resolution, thus there is no need for resampling that reduces the amount of information. Secondly, there are no differences in sensor calibrations or in the spectral resolutions. This makes images visually and spectrally comparable and allows using spectral change detection methods and single signature files after atmospheric corrections (Lu et al. 2004; Franke et al. 2006). Also when provided by USGS Landsat images are often in perfect spatial alignment removing the need for georectification. Altogether single sensor analysis would remove many of errors related to the similarity of the classifications. However the cross-sensor analysis methods need to be developed, because continuous and adequate data from single sensor is often a rarity

(Wulder et al. 2008). Though it is globally important that the continuity of Landsat mission will be guaranteed and data collection and distribution facilities are developed especially in Africa.

If the single sensor data would solve many of these technical aspects related to change detection, manual digitization is able to solve these and also many problems related to classification as such. Manual digitization has multiple benefits as human mind is more able to create cognitive land cover classes than the automated techniques, semantic differences can be minimized with systematic use of image interpretation keys, differences within forests can be recognized from high resolution images, it is easier to georectify high-resolution images and there is no need for resampling to other spatial resolutions (Coppin et al. 2004; Lillesand et al. 2008: 190–250, 485–490). Automated classification techniques are often promoted because they should be time-efficient in large scale, but as relatively small island Unguja could have been easily digitized to few major forest categories during the time that was used to georectify, resample, create the class concepts, interpret unsupervised classifications, draw and modify training areas and post-classify the classifications (Coppin et al. 2004; Lillesand et al. 2008: 545–591). The manual digitization would have also provided quantitatively and spatially significantly more accurate classification and change detection outcomes (Coppin et al. 2004). It might be that in single sensor analysis with Landsat data the automated techniques could be more time-efficient than manual digitization, but these benefits are lost in cross-sensor analysis. Therefore I do not see any reasons why automated classification techniques of medium resolution data should be used instead of manual digitization of high resolution data in medium sized study areas if the required data is available.

Even though single sensor or manually digitized data would be available certain aspects of change would still not be captured. Using land cover classifications from only two different years is not enough to capture the temporality of changes. As mentioned, changes are not always permanent and especially not in the shifting cultivation landscapes (Lambin et al. 2003; Käyhkö & Skånes 2006; Hett et al. 2012). The role of temporal deforestation can be viewed from at least two perspectives. On the other hand it is deforestation as any other form of deforestation. If a certain area used to have forest cover and it is not there now, it has deforested no matter if the cover will return in few years. The other approach relies on FAOs (2000) definition of deforestation, which argues that area needs to be without forest cover over 10 years to be considered deforested. Nevertheless if shifting cultivation fields are considered as deforestation or not, they cause problems in analysis explaining the spatial patterns of the forest change process. The reasons behind spatial distribution of swidden fields are completely different from other deforestation types and cannot therefore be explained by same independent variables (Skole & Tucker 1993; Lambin 1997; Geist & Lambin 2001: 69–71)

Detecting these temporally cyclical changes would have required at least three and preferably four land cover classifications with 10 years interval (Käyhkö & Skånes 2006). The change between first and the second image would be used to determine the recently deforested area. The change between the second and third images would be used to determine the new recently deforested areas and to see has the areas deforested between the first and the second image reforested within the 10 year period and if so they would be considered as temporal deforestation. The same loop would be continued for the change between second-third and third-fourth images, eventually pointing which changes have been temporal and which permanent. These change trajectory analyses have already been used in many similar change detection studies, with good outcomes (Mertens & Lambin 2000; Nagendra et al. 2003; Käyhkö & Skånes 2006; Käyhkö et al. 2011).

7.2.4. Post-analyzing the outcomes of change detection

Using KDE as a tool for mapping the change clusters was originally chosen, because the idea was to map only the deforestation clusters, which would have been impossible with the other clustering methods (Silvermann 1986; Getis & Ord 1992). Later reforestation, forest improvement and degradation were also included and use of other methods would have been possible. Even though of this change, the KDE is still considered as an appropriate tool for determining the clusters of de- and reforestation. However it includes certain inaccuracies. Eventually it is subjective question that how much of the landscape the clusters can cover. In this case 10% of pixels with highest deforestation kernel density value were chosen as clusters, but some other may have chosen this threshold to be 5 or 20%, which would have created completely other outcomes. The tool is also sensitive to the cutoff threshold (Silvermann 1986). There might be a lot of deforested patches in certain area, but if these are relatively far from each other the area does not create a cluster with one kilometer threshold, however it might do so if the threshold is set to 2 or 5 kilometers.

Dividing Unguja to smaller areas based on the soils and protection status provided significant insight to the inner variations of the island. The division was clearly justified as there were clear differences in the change rates and the environmental factors behind the changes (Mertens & Lambin 1997; Serneels & Lambin 2001). The division was quite straightforward to do in Unguja as the divided areas vary from each other so strongly in land uses and the appropriate spatial data was available. However the island could be divided even further for example based on the forest vegetation type, human pressure and land use types. Although it might be extremely difficult to draw borders between these areas. The created function where changes were proportioned against similar changes in entire Unguja was extremely good tool to compress the information of the differences to a single figure.

The created distance analysis showed rather well that the correlation between the used distance variables and deforestation is not always linear (Mertens & Lambin 1997). This helped to interpret the outcomes of the regression models and provided additional knowledge. For example although the vicinity of Zanzibar Town and roads carried no value in regression models, they still showed to influence forest change in distance analysis. The influence may just be non-linear and therefore not detected by the regression methods (Metsämuuronen 2008: 115–119). The distance analysis could have been made better by continuing them even after three kilometers, which may have explained why certain variables had peculiar coefficients in the regression analysis.

The created regression model worked amazingly well, albeit it could be developed significantly further in many ways. As such the model has two major geostatistical deficiencies. Firstly, the statistical significance of regression models cannot be assured if the independent variables are not normally distributed, but this was not tested before the modeling. Variables can be modified to normal distribution with polynomial, squared and logarithmic transformations (Metsämuuronen 2008: 101–103)

Secondly, although the spatial autocorrelation was reduced by making certain that similar cells were at least 1 kilometer apart from each other's, the role of spatial dependency should be better acknowledged both in good and in bad. Congalton (1998) measured that spatial autocorrelation effected land cover of a site until 1,8 kilometers away from it and the Moran's I outcomes from the test site showed that the phenomenon was strongly present even after 1,5 kilometers. Therefore the actions taken to reduce spatial autocorrelation simply may not have been enough. There are few options to handle the spatial autocorrelation in regression modeling, such as spatial error modeling, cellular automata and spatial lag modeling, all having their own pros and cons (Anselin 2002; Verburg et al. 2004). However, only the last is introduced here as possible solution.

Spatial lag model takes the dependent values of the nearby or neighboring cells into account, introduces them to the regression model as independent variables and calculates their regression coefficients. What is not explained by this spatial lag term is then explained by the other variables (Anselin 2002). The regression coefficients could be then used in predictive modeling with or without the spatial lag term. However either using or not using the term has their own downsides. Using it overestimates the deforestation happening near the already deforested cells, as it uses existing state without any knowledge of the previous situation to explain future happenings. It is like estimating that in the future you will have family with two kids as your neighbor based on data from a neighborhood where there is a family with two kids living in 90% of houses. Nevertheless not using the spatial lag term underestimates the spatial autocorrelation and presumes that the status of the neighboring cells will not influence the deforestation happening in the cell in the focus (Anselin 2002; Verburg et al. 2008). The predictive modeling done in this study follows the latter logic and it is therefore better explaining where fully new areas of deforestation will occur than explaining where all of the deforestation is going to happen.

However it is also possible to measure the influence of spatial autocorrelation empirically and use the measured outcomes in predictive modeling. This could be done by having three land cover classification with the same time interval. The change detection outcomes of the two most recent ones would be used in the regression modeling, while the change happened between the first and the second classification would be used to determine where deforestation has happened during this period. When it would be known where deforestation has recently happened in the second image it would be possible to calculate how much deforestation happens between the second and third image near these recently deforested cells.

Besides these geostatistical improvements the modeling could actually be improved with other less technical approaches. The used independent variables are rather good explaining the changes, but they are still far from modeling it perfectly. Therefore new independent variables should be included and the old ones need to be developed. As mentioned the mean NDVI was limited in its abilities and it should be replaced with categorical divisions of forest structures and vegetation types. Also the simple Euclidean distances could be replaced with more advanced accessibility measures as they have been proved better explaining land cover variations (Verburg et al. 2004). Also new sources of accessibility could be introduced. For example vicinity of tourist facilities may explain deforestation happening because spread of tourism. Kernel density of buildings could be replaced with population potential calculations, which not just estimate how much there are buildings within certain threshold, but also takes road accessibility into account and measures how many people can reach certain areas in certain time (Verburg et al. 2004). Also different waterfronts have different meanings as tourism possibly spreads to near sandy beaches or areas with scenic beauty. Also the dependent variable could be transformed from dichotomous to continuous, such as the amount of deforested pixels within a cell. Also it could be disaggregated to original pixel size. The island could also be divided to even smaller sub-areas and deforestation could be explained differently in locations near Zanzibar Town, in the shifting cultivation or the coastal areas.

One of the most intriguing developments would be to test other non-linear regression models. As the distance analyses and other studies have showed the correlations are not always linear and therefore there is a need for non-linear regression models (Mertens & Lambin 1997). One option to add certain non-linearity to the models would be to transform the distance variables with logarithmic and polynomial variations to resemble their line diagrams instead of feeding them in linear. Another method would be to add cutoff distances to variables as the distance analyses have showed that some accessibility measurements like distance to roads only influences until certain threshold. However, even if all these technical and mathematical developments would be done, I assume that the model would still be inadequate, as certain individual or communal level choices or global and regional level changes cannot ever be mathematically modeled, at least not in a single scale of observation.

All the models produce outcomes, but it is another question how accurate these outcomes are. In this study the predictive modeling outcomes were only validated against the cells used to build the model, which creates over positive accuracy estimations (Metsämuuronen 121–123; Mertens & Lambin 2000). The validation could be improved by doing it against the cells not used in the modeling, but also this would create biased outcomes as the dependent variables of the cells not used are influenced by the spatial autocorrelation of the used cells (Congalton 1988). The change process might have also changed after the time period used in the regression modeling and the independent variables may have different influence. Therefore the only solid way to test the usefulness of the predictive model would be to test its outcomes against actual change happened in the future as has been done in multiple similar studies (Mertens & Lambin 1997, 2000; Geoghegan 2001; Serneels & Lambin 2001; Verburg et al. 2002;). This would have required a third image taken after 2009, but cloud free images did not exist.

In ideal situation the regression and predictive modeling would be done with at least 4 images from the same satellite sensor with equal time intervals and perfect spatial alignment (Figure 39). This approach would allow separating temporal deforestation caused for example by shifting cultivation from the permanent deforestation, include empirically measured spatial autocorrelation into the model and allow validation of the outcomes. The first image would be used as a baseline situation, the change between Image 1 and the Image 2 would be used to determine the recently deforested areas in the Image 2, the change happening between Image 2 and Image 3 would be used as the base for the regression modeling and the change between Images 3 and 4 would be used to validate the predictive outcomes of the regression model. Spatial autocorrelation would be empirically measured based on deforestation happening between images 2 and 3 near the recently deforested cells of image 2. Temporal changes such as shifting fields would be excluded from the model with the change trajectory analysis so that areas that have reforested within 10 years interval are not considered as actual deforestation or the modeling could also be mainly focusing on them. The change trajectory analysis would be also used to determine has the change in the time interval used for regression modeling been somehow exceptional and to estimate is the deforestation trend generally increasing or decreasing. If the created model would explain deforestation sufficiently based on the validation, it could be then projected until the future from Image 4.

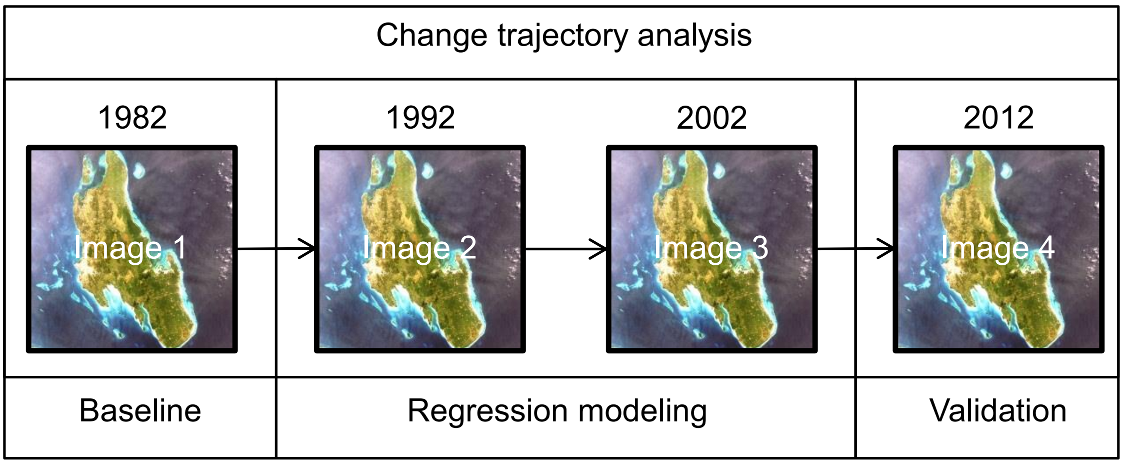


Figure 39. Ideal structure for predictive regression modeling where Image 1 would be used to determine the baseline, change between images 2 and 3 in the actual regression model and the change between images 3 and 4 to validate the predictive modeling outcomes. The change trajectory analysis would be used to measure spatial autocorrelation and to remove short-term changes from permanent deforestation.

7.3. Land Change Science in forest research

It is hard to draw boundaries between Land Change Science, Landscape Ecology and Political Ecology, as there are clear similarities in research topics, questions and methods (Jongman & Pungetti 2004: 2–33; Rindfuss et al. 2004; Turner et al. 2007; Turner & Robbins 2008). Nevertheless what is the used theoretical and methodological framework called, it provided excellent approaches for forest research. Especially change detections based on remotely based land cover classifications have revolutionized the entire fields of land cover and forest research as they provide quantified information about rates, patterns and directions of change (Coppin et al. 2004; Lu et al. 2004; Pontius et al. 2004; Käyhkö & Skånes 2006). However there are always matters that need to be developed. Detecting change as such is an important action, but it is not real-time monitoring. The outcomes are often badly outdated and even though they can be used to create general assumptions where change has happened and where it is probably going to happen, it does not tell where change happens at this moment. Near real-time deforestation monitoring systems have been developed for large forest areas with low spatial resolution data, but these are not adequate for small and diverse surroundings, such as Zanzibar, because of their pixel size (Anderson et al. 2005). Hopefully in the future as the satellite sensors develop and data sharing policies open up there will be medium or high resolution data with high temporal coverage allowing continuous monitoring of the decreasing forest resources.

There is also a certain methodological gap between detecting changes and connecting them to the proximate and underlying causes. Change can be detected in rather detailed accuracy, regional variations can be emphasized and happened changes can be easily connected to environmental factors, but eventually these analyses do not provide any empirical evidence about the causes behind the changes (Veldkamp & Lambin 2001; Verburg et al. 2004). Certainly cross-sectional sub-area analysis may show that there are change differences between two areas and spatial regression analysis may explain patterns of change with environmental factors, but differences between areas can be caused by thousands of different causes, which cannot be separated and also the environmental factors and causes are so tangled that it is impossible to link the changes in regression coefficients to actual causes. (Rindfuss et al. 2004; Verburg et al. 2004) More aspatial regression analysis using administrative level socio-economic statistics to explain deforestation in large study areas may show that changes in certain societal aspect cause changes also in forests, but even then the happened changes cannot be undressed from interdependencies of local context (Tole 2001; Verburg et al. 2004; Aguiar 2007). Many have relied connecting causes to process with narrative analysis based on secondary documents or expert and local knowledge (Geist & Lambin 2001: 18; Verburg et al. 2004). Although these approaches probably provide the most reliable outcomes, I see them vulnerable to simplifications and repetitions of well-established “truths”.

Land Change Science relies heavily on the positivist tradition and empirically measured observations. It is therefore eager looking explanations to changes with quantifiable data that can be fed into statistical modeling methods (Veldkamp & Lambin 2001; Rindfuss et al. 2004; Verburg et al. 2004; Turner et al. 2007). However the causes behind changes are so interlinked and multi-scaled that they cannot ever be separated or simultaneously understood at multiple scales only with the quantitative methods. Therefore I see it extremely important not to get stuck with single predefined frameworks, such as LCS, but to be pragmatically experimental. What I mean with this gibberish is more daring unification of measuring quantitative and explaining qualitative approaches. This could be done for example by using maps of empirically measured changes to stimulate the interviews of stakeholders to go beyond simplifications, by asking experts to connect environmental factors to proximate and underlying causes simultaneously at multiple scales and by connecting local uses to forest areas with PGIS methods (Fagerholm & Käyhkö 2009; Fagerholm et al. 2011, 2012).

Eventually though, the question is not only about how well forest changes can be understood scientifically, but also about the usability of this knowledge. From my opinion the methods used here are unable to create the full benefits for the communities. Historical change detections do not please the needs of real-time monitoring and the statistical methods often only emphasize the already known facts, while true benefits are only achieved if new information, knowledge or even wisdom could be produced.

8. Conclusions

- At the moment there is relatively lot of forests cover in Unguja (572 km^2 & 38%), but majority of it is in the low-lying coral rag scrublands and agroforests.

- The forest cover is declining rapidly (0,82–1,18% per year) and in 2025 there is going to be only 485 km^2 of forests left. Besides deforestation the forests are also suffering from vertical degradation and spatial changes.

- Deforestation is most severe in the communal indigenous coral rag forests, but also agroforests are decreasing. The government protected areas on the other hand are able to increase their forest cover.

- More precisely deforestation concentrates to areas near the coastline, population and Zanzibar Town, while soil, elevation, vegetation and road distance differences are not explaining the spatial patterns of the process.

- The deforestation is mainly driven by small holder farmers, but also urban dwellers, administration, rural communities and tourism actors are involved. Shifting cultivation, spreading of more permanent agriculture, urban sprawl, village growth, tourism developments, planned resettlement actions and logging are the main proximate causes. The proximate causes are linked to such underlying drivers as population growth, in-migration, urbanization, tourism development, economic growth and land tenure

- Semantic concepts, their technical executions, georectification, resampling, spatial and spectral resolution differences caused substantial biases in the estimations. Many of these errors could have been solved by using only single sensor data.

- The theoretical and methodological frameworks of LCS and deforestation research provide many usable tools for forest change research enabling detecting main areas of change, differences between areas and connecting changes to environmental causes. However these tools are limited in providing empirical evidence about the proximate and underlying causes of change.

Acknowledgements

This thesis is part of the research project *Changing land use and forest management practices and multidimensional adaptation strategies in Zanzibar Tanzania* (2010-2013), which is funded by the Academy of Finland and the Universities of Turku and Dar es Salaam. I owe thanks to the project, Turku Geographical Society, Uskelan Opintorahastosäätiö and University of Turku for funding my research and field working in Zanzibar.

I would like to thank the Department of Forestry and Non-Renewable Resources under the Government of Zanzibar for their support. Especially I wish to express my gratitude to forest officers MSc Miza Khamis and BSc Abbas Mzee for all their help in the field. Without you I would have been lost, literally. Thanks belong also to Sustainable Management of Land and Environment –project and project staff for sharing data that was essential for this research. Special thanks to my supervisor Adjunct Prof. Niina Käyhkö for all the encouragement when thinks have not gone so well and for pushing me towards more critical thinking. Thanks to Dr. Nora Fagerholm, MSc Timo Pitkänen, MSc Salla Eilola and MA Zahor K. Zahor for all your help. Finally I would like to thank Abubakar Saleh Moh'd for all the good times wilding and fishing in Zanzibar. Pumzika Kwa Amani.

References:

Aguiar, A.P., G. Câmara & M.I. Escada (2007). Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecological Modelling* 209, 169–188.

Ahlqvist, O. (2004). A Parameterized Representation of Uncertain Conceptual Spaces. *Transactions in GIS*, 8, pp. 493–514.

Ahlqvist, O. (2008). In search of classification that supports the dynamics of science: The FAO Land Cover Classification System and proposed modifications. *Environment and Planning B: Planning and Design,* 35. 169–186.

Ahrends, A., N.D. Burgess, S. Milledge, M. Bulling, B. Fisher, J. Smart, G.P. Clarke, B. Mhoro & S- Lewis (2010). Predictable waves of sequential forest degradation and biodiversity loss spreading from an African city. *PNAS* 107:33, pp. 14556–14561.

Alcamo, J., K. Kok, G. Busch, J. Priess, B. Eickhout, M. Rounsevell, D. Rothman & M. Heistermann (2006). Searching for the future of land: Scenarios from the local and global scale. In: Lambin, E. & H. Geist (eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts.* 117–135. Springer-Verlag Berlin Heidelberg, Germany.

Alcorn, J.B. (1993). Indigenous peoples and conservation. *Conservation Biology* 7, 424–426.

Ameyibor, S., T. Basteck, C. Bierbaum, N. Frommeld, N. Giaourakis, K. Hackenbroch, A. Kirchberg, A. Kutsch, M. Mendel, S. Schlichting, S. Schmelz, J. Schoppengerd, A. Schulitz, J. Schützendorf, C. Sufryd & T. Vorberg (2003). Informal Settlements Development on Zanzibar: A Study on the Community Based Provision of Storm Water Management. *SPRING Working Papers No 21,* pp, 134. SPRING Centre University of Dortmund, Dortmund.

Andersson, L., Y.E. Shimabukuro, R. DeFries & D. Morton (2005). Assessment of deforestation in near real time over the Brazilian Amazon using multitemporal fraction images derived from Terra MODIS. *Geoscience and Remote Sensing Letters*, 2:3, 315–318.

Angelesn, A. & D. Kaimowitz (1999). Rethinking the causes of deforestation: Lessons from economic models. *The World Bank Research Observer* 14:1. pp. 73–98.

Angelsen, A. (2007). Forest cover change in space and time: Combining von Thünen and Forest Transition theories. *World Bank Policy Research Working Paper* 4117, pp. 1–43.

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis*, 27:2, 93–115.

Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, 27, 247–267.

Balick, M.J. & R. Mendelsohn (1992). Assessing the economic value of traditional medicines from tropical rain forests. *Conservation Biology* 6, 128–130.

Barbier, E., J. Burgess & A Grainger (2010). The Forest Transition: Towards a More Comprehensive Theoretical Framework. *Land Use Policy* 27, pp. 98–107.

Benz, U., P. Hofmann, G. Willhauck, I. Lingenfelder & M. Heynen (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, pp. 239– 258

Brinkmann, K., A. Patzelt, E. Schlecht & A. Buerkert (2011). Use of environmental predictors for vegetation mapping in semi-arid mountain rangelands and the determination of conservation hotspots. *Applied Vegetation Science*, 14, 17–30.

Brown, S. & A.E. Lugo (1990). Tropical Secondary Forests. *Journal of Tropical Ecology,* 6: 1. 1–32.

Burgess, N.D., G.P. Clarke, J. Madgwick, S.A. Robertson & A. Dickinson (2000a). Status and vegetation; distribution and status. In N.D. Burgess & G.P. Clarke (eds.), *Coastal Forests of Eastern Africa*. IUCN Forest Conservation Programme, Gland and Cambridge. Pp. 71–81.

Burgess, N.D., D. Kock, A. Cockle, C. FitzGibbon, P. Jenkins & P. Honess (2000b). Biodiversity Values; mammals. In N.D. Burgess & G.P. Clarke (eds.), *Coastal Forests of Eastern Africa*. IUCN Forest Conservation Programme, Gland and Cambridge. Pp. 173–190.

Chomitz, K.M. & D.A. Gray (1996). Roads, land use and deforestation: A spatial model applied to Belize. *The World Bank Economic Review* 10: 3, 487–512.

Clark, J. & T. Bobbe (2007). Using remote sensing to map and monitor fire damage in forest ecosystems. In M. Wulder & S. Franklin (eds.) *Understanding Forest Disturbance and Spatial Patter: Remote Sensing and GIS Appproaches*. pp. 113–131. Taylor & Francis Group, New York.

Congalton, R. (1988). A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering & Remote Sensing,* 54:5. 593–600.

Contreras-Hermosilla, A. (2000). The Underlying Causes of Forest Decline. *CIFOR Occasional Paper* 30. pp. 1–25.

Dallu, A.I.M. (2004). *The coastal forests of Tanzania: a national synthesis report for the preparation of WWF East African Coastal Forest Ecoregion Programme*. Forestry and Beekeeping Division, Ministry of Natural Resources and Tourism, Dar es Salaam, 29 pp.

D'Almeida, C., C.J. Vörösmarty, G.C. Hurtt, J.A. Marengo, S. Lawrence Dingman & B.D. Keim (2007). The effects of deforestation on the hydrological cycle in Amazon: a review on scale and resolution. *International Journal of Climatology* 27, pp. 633–647.

Department of Comercial Crops, Food and Forestry (DCCFF) (2008). *Zanzibar National Forest Resources Management Plan 2008 – 2020.* 51 p. DCCFF, Zanzibar Town.

DeFries, R., T. Rudel, M. Uriarte & M. Hansen (2010). Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience* 3, 178–181

Di Gregorio, A. (2005). *Land cover classification system: Classification concepts and user manual. Software version 2. p. 190. Food and agriculture organization of the United Nations (FAO), Rome*

ESRI (2012). Modeling spatial relationship: Selecting a fixed distance band value. 12.11.2012. <http://help.arcgis.com/en/arcgisdesktop/10.0/help/#/Modeling_spatial_relationships/005p00000005000000/>

Fagerholm, N & N. Käyhkö (2009). Participatory mapping and geographical patterns of the social landscape values of rural communities in Zanzibar, Tanzania. *Fennia* 187:1. pp. 43–60.

Fagerholm, N. (2012). Community inhabitants values and benefits in dynamic tropical forest landscapes: Participation and spatial analysis in landscape knowledge integration. *Annales Universitatis Turkuensis AII 272.* Painosalama Oy, Turku.

Fagerholm, N., N. Käyhkö, F. Ndumbaro, M. Khamiz (2012). Community stakeholders' knowledge in landscape assessments – Mapping indicators for landscape services. *Ecological Indicators* 18, pp. 421–433.

FAO (Food and Agriculture Organization) (1995). *Forest resource assessment 1990. Global synthesis*. 44 p. FAO, Rome.

FAO (2000). *On Definition of Forest and Forest Change.* p. 15. Forest Resource Assessment Program, Working paper 33, Rome.

FAO (2010). *Global Forest Resources Assessment 2010: Main report*. p. 378. Food and Agricultural Organization of the United Nations, Rome.

Fisher, B. (2010). African exception to drivers of deforestation. *Natural Geoscience* 3, 375–376.

Foody, G. & R. Hill (1996). Classification of tropical forest classes from Landsat TM data. *International Journal of Remote Sensing*, 17:2, 2353–2367.

Foody, G. (2004). Sub-pixel methods in remote sensing. In: De Jong, S. & F. Van der Meer (eds.): *Remote sensing image analysis: Including the spatial domain*. 37–49. Springer, Neatherlands.

Foster, A & M. Rosenzweig (2003). Economic growth and the rise of forests. *The Quarterly Journal of Economics* 118, pp. 601–637.

Fox, J., T. Rambo, D. Donovan, L.T. Cuc, T. Giambelluca, A. Ziegler, D. Plondke, T.D. Vien, S. Leisz & D.M. Truong (2003). Linking household and remotely sensed data for understanding forest fragmentation in Northern Vietnam. In: Fox, J., R. Rindfuss, S. Walsh & V. Mishra (eds.): *People and the environment: Approaches for linking household and community surveys to remote sensing and GIS,* 201–221. Kluwer Academic Publishers, Norwell.

Franke, J. V. Heinzel & G. Menz (2006). Assessment of NDVI differences caused by sensor-specific relative spectral response functions. *Proceeding of the International Geoscience and Remote Sensing Symposia. Denver, USA. July 31st – August 4th 2006.*

Garcia, M, R. Riano, E. Chuvieco & FM. Danson (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, 114. 816–830.

Geist, H. & E. Lambin (2001). *What drives tropical deforestation: A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence.* LUCC Report Series No. 4. 136 p.

Geist, H., W. McConnell, E. Lambin, E. Moran, D. Alves & T. Rudel (2006). Cause and trajectories of land-use/cover change. In: Lambin, E. & H. Geist (eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts.* 41–70. Springer-Verlag Berlin Heidelberg, Germany.

Geoghegan, J., S. Villar, P. Klepeis, P. Mendoza & Y. Ogneva-Himmelberger (2001). Modeling tropical deforestation in the southern Yucatán peninsular region: Comparing survey and satellite data. *Agriculture, Ecosystems and Environment* 85, pp. 25–46.

Getis, A. & J.K. Ord (1992). The analyses of spatial association by use of distance statistics. *Geographical Analysis*, 24:3, 189–206.

Godoy, F., K. Tabor, N.D. Burgess, B. Mbilinyi, J. Kashaigili & M. Steininger (2011). Deforestation and CO^2 emission in coastal Tanzania from 1990 to 2007. *Environmental Conservation* 39: 1, pp. 62–71.

Grove, M. (2011). A spatio-temporal kernel method for mapping changes in prehistoric land-use patterns. *Archaeometry*, 53:5, 1012–1030.

Gössling, S. (2001). Tourism, economic transition and ecosystem degradation: interacting processes in a Tanzanian coastal community. *Tourism Geographies,* 3, pp. 430–453.

Hansen, M. C., R. S. Defries, J. R. G. Townshend & R. Sohlberg (2000): Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21:6-7, pp.1331-1364.

Hardin, G. (1968). The tragedy of commons. *Science,* 162: 3859, pp. 1243–1248.

Healey, S., W. Cohen, Y. Zhiqiang & R. Kennedy (2007). Remotely sensed data in the mapping of forest harvest patterns. In M. Wulder & S. Franklin (eds.) *Understanding Forest Disturbance and Spatial Patter: Remote Sensing and GIS Appproaches*. pp. 63–85. Taylor & Francis Group, New York.

Hett, C., J.C. Castella, A. Heinimann, P. Messerli, J.-L. Pfund (2012). A landscape mosaics approach for characterizing swidden systems from a REDD+ perspective. *Applied Geography*, 32:2, pp. 608–618.

Hettige, M.L. (1990). *Land evaluation and land suitability classification – Unguja & Pemba Island.* 142 pp. Food and Agriculture Organization of the United Nations.

Houghton, R., D. Skole, C. Nobre, J. Hackler, K. Lawrence & W. Chementowski (2000). Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. *Nature* 403, 301–304.

Irwin, E. & J. Geoghegan (2001). Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems and Environment* 85, pp. 7–23.

Islam, K.R. & R.R. Weil (2000). Land use effects on soil quality in a tropical forest ecosystem of Bangladesh. *Agriculture, Ecosystem and Environment* 79, pp. 9–16.

Jongman, R. & G. Pungetti (2004). *Ecological Networks and Greenways: Concept, Desing and Implementation.* 345 p. Cambridge University Press, Cambridge.

Kaimowitz, D. & A. Angelsen (1998). *Economic Models of Tropical Deforestation. A Review,* 153 p. Centre for International Forestry Research, Bogor.

Khan, S. & S. Khan (2009). Assessing poverty–deforestation links: Evidence from Swat, Pakistan. *Ecological Economics*, 68, pp. 2607–2618

Klein, R. (2008). The epadphic landscape of Unguja island, Zanzibar: An exploratory study on the relation between soil variability and the landscape. *Turku University Department of Geography Publications B Nr. 11*, p. 94. Digipaino Oy, Turku.

Klein, R. & J. Käyhkö (2008). The general character of Unguja landscape. In: Käyhkö, N., M. Orjala & J. Mustelin (eds.): Sustainable landscapes in Zanzibar, 33–38. Turku University Department of Geography Publications A: 172. Digipaino, Turku. X

Kok, K. & A. Veldkamp (2001). Evaluating impact of spatial scales on land use pattern analysis in Central America. *Agriculture, Ecosystems and Environment* 85, pp. 205–221.

Kombo, Y.H. & M.M. Kitwana (1997). *Coral rag soil fertility improvement. Farmers' indigenous knowledge and practices.* Commission of Natural Resources, Zanzibar. Forestry Technical Paperno. 72. 12 pp.

Krain, E. (1998). *The agrarian constitutions of Zanzibar and its impact on agricultural development*. Studien für Wirtschaft- und Agrarpolitik B 17. M. Wehle, Witterschlick & Bonn. 206 pp.

Käyhkö, N, N. Fagerholm, B.S. Asseid & A.J. Mzee (2011). Dynamic land use and land cover changes and their effect on forest resources in a coastal village of Matemwe, Zanzibar, Tanzania. *Land Use Policy* 28: 1, 26–37.

Käyhkö, N & H: Skånes (2006). Change trajectories and key biotopes—Assessing landscape dynamics and sustainability. *Landscape and Urban Planning,* 75, pp. 300–321.

Lambin, E. (1997). Modelling and monitoring land-cover change process in tropical regions. *Progress in Physical Geography*, 21: 3, pp. 375–393.

Lambin, E., B.L. Turner, H. Geist, S. Agbola. A. Angelsen, J. Bruce, O. Coomes, R. Dirzo, G. Fischer, C. Folke, P.S. George, K. Homewood, J. Imbernon, R. Leemans, X. Li, E. Moran, M. Mortimore, P.S. Ramakrishnan, J. Richards, H. Skånes, W. Steffen, G. Stone, U. Svedin, T. Veldkamp, C. Vogel & J. Xu (2001). The cause of land-use and land-cover change: moving beyond the myths. *Global Environmental Change* 11. pp. 261–269.

Lambin, E., H. Geist & E. Lepers (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, 28, pp. 205–241.

Lambin, E., H. Geist & R. Rindfuss (2006). Introduction: Local Processes with Global Impacts. In. Lambin. E.F. & H. Geist (Eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts, 41–70.* Springer, Germany.

Laurance, W.F. (1999). Reflections on the tropical deforestation crisis. *Biological Conservation* 91, 109–117.

Lillesand, T., R. Kiefer & J. Chipman (2008). *Remote Sensing and Image Interpretation.* 6th ed._p. 757. Library of Congress Cataloging-in-Publication Data, U.S.A.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang & J. W. Merchant (2000): Development of a global land cover characteristics database and

IGBP DISCover from 1 km AVHRR data. _International Journal of Remote Sensing_, 21:6-7, pp. 1303-1330.

Lu, D., P. Mausel, E. Brondizio & E. Moran (2004). Change detection techniques. _International Journal of Remote Sensing_, 25:12, pp. 2365–2407.

Ludeke, A., R. Maggio, L. Reid (1990). An analysis of anthropogenic deforestation using logistic regression and GIS. _Journal of Environmental Management_ 31, pp. 247–259.

Lunetta, R. & J. Lyon (2000). _Remote Sensing and GIS Accuracy Assessments._ 304 p. CRC Press LLC, Boca Raton

Maass, J., P. Balvanera, A. Castillo, G. C. Daily, H. A. Mooney, P. Ehrlich, M. Quesada, A. Miranda, V. J. Jaramillo, F. García-Oliva, A. Martínez-Yrizar, H. Cotler, J. López-Blanco, A. Pérez-Jiménez, A. Búrquez, C. Tinoco, G. Ceballos, L. Barraza, R. Ayala, and J. Sarukhán (2005). Ecosystem services of tropical dry forests: insights from long-term ecological and social research on the Pacific Coast of Mexico. _Ecology and Society_ 10:1, pp. 1–23.

Maeda, E. E. (2011). _Agricultural expansion and climate change in the Taita Hills, Kenya: An assessment of potential environmental impacts_. 58 p. Helsinki University Print, Helsinki.

Maconachie, R., A. Tanko & M. Zakariya (2009). Desending the energy ladder? Oil price shocks and domestic fuel choices in Kano, Nigeria. _Land Use Policy_, 26:4, 1090–1099.

Martello, R., P. Dargusch & Medrilizam (2010). A Systems Analysis of Factors Affecting leakage in Reduced Emission From Deforestation and Degradation Project in Tropical Forests in Developing Nations. _Small-scale Forestry_ 9, pp. 501–516.

Masoud, R.S. (1991). Fuelwood Use in Zanzibar Town. _Zanzibar Forestry Development Project Technical Paper Number 3._ 1–51.

Mather, P. (2005). _Computer Processing of Remotely Sensed Images._ 3rd edition. 312 p. John Wiley & Sons, England.

McConnell, W.J. & E. Keys (2005). Meta-analysis of agricultural change. In: Moran, E.F. & E. Ostrom (eds.): _Seeing the forest and the trees: Human-environment interactions in forest ecosystems_, 325–353.. MIT Press, Cambridge London.

McPeak, J. & C. Barrett (2001). Differential risk exposure and stochastic poverty traps among East African pastoralists. _American Journal of Agricultural Economics_ 83, pp. 674–679

Mertens, B. & E. Lambin (1997). Spatial modeling of deforestation in southern Cameroon: Spatial disaggregation of diverse deforestation process. _Applied Geography_ 17: 2, 142–162.

Mertens, B., W. Sunderlin, O. Ndoye & E. Lambin (2000). Impact of macroeconomic change on deforestation in South Cameroon: Integration of household survey and remotely-sensed data. _World Development_ 28: 6, 983–999.

Meyfroidt, P & E. Lambin (2009). Forest transition in Vietnam and displacement of deforestation abroad. PNAS 106: 38. pp. 16139–16144.

Milledge, S., I. Gelvas & A. Ahrends (2007). *Forestry, Governance, National Development. Lessons Learned from Logging Boom in Southern Tanzania*. TRAFFIC East/Southern Africa / Tanzania Development Partners Group / Ministry of Natural Resources of Tourism, Dar es Salaam, Tanzania. 252 p.

Mustelin, J. (2008). Tourism's effects on the coastal communities in the east coast of Unguja island. In Käyhkö, N., M. Orjala, J. Mustelin (eds.): *Sustainable landscapes in Zanzibar*, 105–111. Turku University Department of Geography Publications A 172

Nagendra, H, J. Southworth & C.M. Tucker (2003). Accessibility as a determinant of landscape transformation in western Honduras: linking pattern and process. *Landscape Ecology* 18: 2, 141–158.

Nagendra, H., D. K. Munroe & J. Southworth (2004). From pattern to process: Landscape fragmentation and the analysis of land use/land cover change. *Agriculture, Ecosystems and Environment* 101, 111–115.

Nagendra, H. (2007). Drivers of reforestation in human-dominated forests. *Proceedings of the National Academy of Science of the United States of America* 104: 39, 15218–15223.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment, 80.* 88–99.

NBS (National Bureau of Statistics) (2004a). *2002 Population and Housing Census: Volume IV: District Profile: Central*. 130 p. Central Census Office, Dar es Salaam.

NBS (2000b). *2002 Population and Housing Census: Volume IV: District Profile: North "A"*. 132 p. Central Census Office, Dar es Salaam.

NBS (2000c). *2002 Population and Housing Census: Volume IV: District Profile: North "B"*. 133 p. Central Census Office, Dar es Salaam.

NBS (2000d). *2002 Population and Housing Census: Volume IV: District Profile: South.* 130 p. Central Census Office, Dar es Salaam.

NBS (2000e). *2002 Population and Housing Census: Volume IV: District Profile: Urban*. 112 p. Central Census Office, Dar es Salaam.

NBS (2000f). *2002 Population and Housing Census: Volume IV: District Profile: West.* 130 p. Central Census Office, Dar es Salaam.

NBS (2005). *Tanzania demographic and health survey 2004–2005: Key findings for Zanzibar.* 19p. Calverton, USA.

OCGS (Office of Chief Government Statistician) (2006). *Zanzibar Sample Census of Agriculture 2002/2003: Small Holder Agriculture: VII Crop Section.* 247 p. Zanzibar Town, tanzania

OCGS (2008). *Zanzibar statistical abstract 2007*. 85 p. Zanzibar Town, Tanzania.

OCGS (2010). *Zanzibar Socio– Economic Survey Report.* 93 p. Zanzibar Town, Tanzania

Olander, L., H. Gibbs, M. Steininger, J. Swenson & B. Murray (2008). Reference scenarios for deforestation and forest degradation in support of REDD: a review of data and methods. *Environmental Research Letters* 3:2, 1–11.

Olsson, H. (1995). Reflectance calibration of Thematic Mapper data for forest change detection. *International Journal of Remote Sensing*, 50, pp. 221–230.

Omotola, S. (2006). The Next Gulf? Oil Politics, Environmental Apocalypse and Rising Tension in the Niger Delta, *ACCORD Occasional Paper Series* 1: 3, 1–31.

Ord, K. & A. Getis (1996). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, 27:4, 286–306.

Orjala, M. (2006). Lifeworld and its reflections in rural communities: The interaction between an individual, a village community and nature on the island of Unguja in Zanzibar. *Turku University Department of Geography Publications B*:8, 124 p. Digipaino, Turku.

Prins, E. & G.P. Clarke (2007). Discovery and enumeration of Swahilian Coastal Forests in Lindi region, Tanzania, using Landsat TM data analysis. *Biodiversity Conservation* 16, pp. 1551–1565.

Puyravaud, J-P. (2003). Standardizing the calculation of the annual rate of deforestation. *Forest Ecology and Management*, 177, 593–596.

RGS (Revolutionary Government of Zanzibar) (2004). *State of the environment report for Zanzibar,* 36 p. Zanzibar Town, Tanzania.

RGZ (2006). *2004/2005 Household Budget Survey*. 179 p. Zanzibar Town, Tanzania.

RGZ (2007). *Zanzibar strategy for growth and reduction of poverty.* 107 p. Zanzibar Town, Tanzania.

RGZ (2009). *Zanzibar human development report 2009: Towards pro poor growth.* 150 p. Zanzibar Town, Tanzania.

Rindfuss, R., S. Walsh, B.L. Turner II, J. Fox & V. Mishra (2004). Developing a science of land change: Challenges and methodological issues. *PNAS* 101:39, pp. 13976–13981.

Rogan, J. & J. Miller (2007). Integrating GIS and remotely sensed data for mapping forest disturbance and change. In M. Wulder & S. Franklin (eds.) *Understanding Forest Disturbance and Spatial Patter: Remote Sensing and GIS Appproaches*. pp. 133–171. Taylor & Francis Group, New York.

Rosero-Bixby, L. & A.Palloni (1996) Population and deforestation in Costa Rica. *CDE Working Paper No. 96-19*.

Rudel, T.K. & J. Roper (1997). Forest fragmentation in the humid tropics: A cross-national analysis. *Singapore Journal of Tropical Geography* 18: 1, 99–109.

Rudel, T.K. (2005). *Tropical forests: Regional paths of destruction and regeneration in the late 20th century*, 239 p. Columbia University Press, New York.

Rudel, T.K., O. Coomes, E. Moran, F. Achard, A. Angelsen, J. Xu & E. Lambin. (2005). Forest transitions: towards a global understanding of land use change. *Global Environmental Change* 15, pp. 23–31.

Rudel, T., R. DeFries, G. Asner & W. Laurance (2009). Changing drivers of deforestation and new opportunities for conservation. Conservation Ecology, 29:6, 1396–1405.

Sahin, V. & M.J. Hall (1996). The effects of afforestation and deforestation on water yields. *Journal of Hydrology* 178. pp. 293–309.

Scherr, S (2000). A downward spiral? Research evidence on the relationship between poverty and natural resource degradation. *Food Policy* 25:4, pp. 479–498.

Seaman, D. & R. Powell (1996). An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology*, 77:7, 2075–2085.

Serneels, S. & E. Lambin (2001). Proximate causes of land-use change in Narok District, Kenya: a spatial statistical model. *Agriculture, Ecosystem and Environment* 85, 65–81.

Serneels, S. M. Linderman & E. Lambin (2007). A Multilevel Analysis of the Impact of Land Use on Interannual Land-Cover Change in East Africa. *Ecosystems* 10, 402 – 418.

Siex, K. (2011). *Protected area spatial planning for Unguja and Pemba island, Zanzibar.* p. 42 Wildlife Conservation Society (WCS) Africa Program, Bronx.

Silvermann, B.W. (1986). *Density estimation for statistics and data analysis.* 22 p. Chapman and Hall, London.

Sitari, T. (2005). Forestry in the community lideworld in Unguja island Zanzibar. In: Sitari, T. (eds). *Forestry, Community and Biodiversity in Zanzibar,* 3. Turku University Department of Geography Publications B, pp. 37 – 58.

Skole, D. & C. Tucker (1993). Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science* 260, 1905–1910.

Southworth, J. H. Nagendra & C.M. Tucker (2002). **F**ragmentation of a Landscape: incorporating landscape metrics into satellite analyses of land-cover change. *Landscape Research* 27: 3*,* pp. 253–269.

Souza, C.M. & P. Barreto (2000). An alternative approach for detecting and monitoring selectively logged forests in the Amazon. *International Journal of Remote Sensing*, 21, 173–179.

Stehman, S. & R. Czaplewski (1998). Desing and Analysis of Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing of Environment,* 64. 331–344.

Swetnam, R.D., B. Fisher, B.P. Mbilinyi, P.K.T. Munishi, S. Willcock, T. Ricketts, S. Mwakalila, A. Balmford, N.D. Burgess, A.R. Marshall & S.L. Lewis (2011). Mapping socio-economic scenarios of land cover change: A GIS method to enable ecosystem service modeling. *Journal of Environmental Management* 92, pp. 563–574.

Swinton, S.M, G. Escobar & T. Reardon (2003). Poverty and environment in Latin America: concepts, evidence and policy implications. *World Development* 31:11, 1865–1872.

Tabor, K., N.D. Burgess, B. Mbilinyi, J. Kashaigili & M. Steininger (2010). Forest and woodland cover and change in coastal Tanzania and Kenya, 1990 to 2000. *Journal of East African Natural History,* 99:1, pp. 19–45.

TFCG (Tanzanian Forest Conservation Group) (2006). Coastal Forests of Kenya and Tanzania. 19.11.2012. <http://coastalforests.tfcg.org/index.html>

The Royal Norwegian Embassy (2010). Norway supports major climate change initiative in Zanzibar. 20.11.2012. *<http://www.norway.go.tz/News_and_events/News/Embassy-Press-Releases/Norway-supports-major-climate-change-initiative-in-Zanzibar/>*

Turner, B. L., E. Lambin & A. Reenberg (2007). The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 104: 52, 20666–20671.

Turner, B.L., & P. Robbins (2008). Land-Change Science and Political Ecology: Similarities, Differences, and Implications for Sustainability Science. *Annu. Rev. Environ. Resour*, 33, 295–316

Törhönen, M. (1998). *A thousand and one nights of land tenure. The past, present and future of land tenure in Zanzibar*. 99 pp. Common Estate-Series. The Royal Institution of Chartered Surveyors.

Unruh, J., H. Nagendra, G.M. Green, W.J. McConnell, N. Vogt (2005) Cross-continental comparisons: Africa and Asia. In: Moran, E.F. & E. Ostrom (eds*): Seeing the forest and the trees: Human-environment interactions in forest ecosystems*, 303–324. MIT Press, Cambridge London.

Veldkamp A. & E. Lambin (2001). Predicting land use change. *Agriculture, Ecosystem and Environment* 85. pp. 1–6.

Verburg, P., W. Soepboer, A. Veldkamp, R. Limpiada, V. Espaldon & S. Mastura (2002). Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model. *Environmental Management* 30: 3, pp. 391–405.

Verburg, P. & A. Veldkamp (2003). Projecting land use transitions at forest fringes in the Philippines at two spatial scales. *Landscape Ecology,* 19, pp. 77–98.

Verburg, P., P. Schot, M. Dijst & A. Veldkamp (2004). Land use change modelling: Curent practice and research priorities. *GeoJournal* 61, pp. 309–324.

Verburg, P., K. Overmars & N. Witte (2004b). Accessibility and Land-Use Patterns at the Forest Fringe in the Northeastern Part of the Philippines. *The Geographical Journal*, 170:3, 238-255.

Verburg, P., K. Kok, R. Pontius & A. Veldkamp (2006). Modelling land-use and land-cover change. In: Lambin, E. & H. Geist (eds.): *Land-Use and Land-Cover Change: Local Processes and Global Impacts.* 117–135. Springer-Verlag Berlin Heidelberg, Germany.

Walker, R. (2004). Theorizing land-cover and land-use change: The case of tropical deforestation. *International Regional Science Review* 27: 3, pp. 247–270.

Weiers, S., G. Groom & M. Wissen (2002). Comparability and subjectivity of land cover maps produced with digital image classification techniques: some recent experiences from Denmark and northern Germany. *Danish Journal of Geography, 102.* 59–77.

Wilson, E. O (1988). The current state of biodiversity. In: Wilson, E.O. & F. M. Peters (eds): *Biodiversity,* pp 3–20. National Academy Press, Washington D.C.

Wright, S.J, & H. Muller-Landau (2006). The Future of Tropical Forest Species. *Biotropica*, 38:3, 287–301.

Wulder, M, C. Butson & J. White (2008). Cross-sensor change detection over a forested landscape: Options to enable continuity of medium spatial resolution measures. *Remote Sensing of Environment* 112, 23–31.

WWF Tanzania Country Office (2012). *Synthesis Baseline Report for Coastal Forests in Tanzania.* p. 174. WWF-Tanzania, Dar es Salaam.

APPENDIX 1. Field observation sheet (Niina Käyhkö & Markus Kukkonen 2011)

ZANZIBAR –PROJECT 2011, FIELD OBSERVATION SHEET

Site ID: ___________ Observer(s): _____________ Date: _______________

Photographs:______________

Place (shehia, village, site): ____

GPS X: GPS Y: GPS Z:

Land cover type/class (describe):

Topography within the site:(1) flat (2) slope (inclination N S E W) (3) undulating

Soil type: (1) coral rag (> 75%) (2) semi-coral (~50%) (3) loose soils dominating (4) sand (5) other

Moisture conditions: |---|

(circle along the gradient) **dry** **wet**

Other edaphic (bedrock/soil/topography/) observations of the site:

Ground layer coverage: (1) bare ground ___% (2) grasses and herbs ___% (3) cultivated ____%

Number of tree/scrub layers: 1 2 3 4

Layer heights: (1) <1 m (2) 1-3 m (3) 3-5 m (4) 5-10 m (5) >10 m

Dominating tree canopy shape: (1) wide/round (2) narrow (3) mixed

Tree/scrub coverage: (1) open (<25%) (2) semi-open (25-75%) (3) closed (>75%)

Spatial pattern: (1) (2) (3) (4)

Dominant tree/scrub species (list):

Ongoing/annual land use activities within the area?

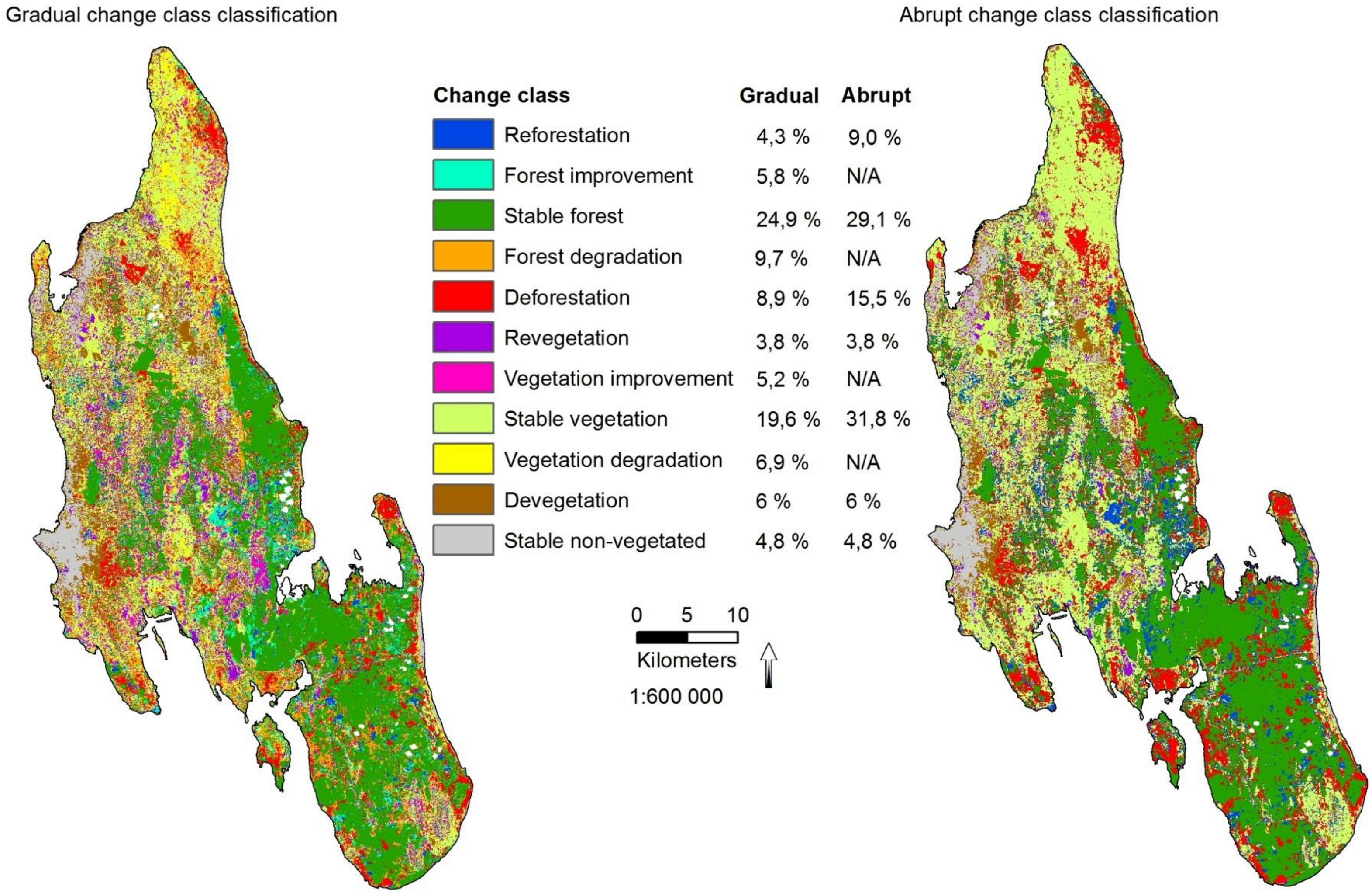
Visible evidences of past/historical land uses? (e.g. planting of trees, extraction of coral/sand etc.)

APPENDIX 2. Class-by-class change statistics:

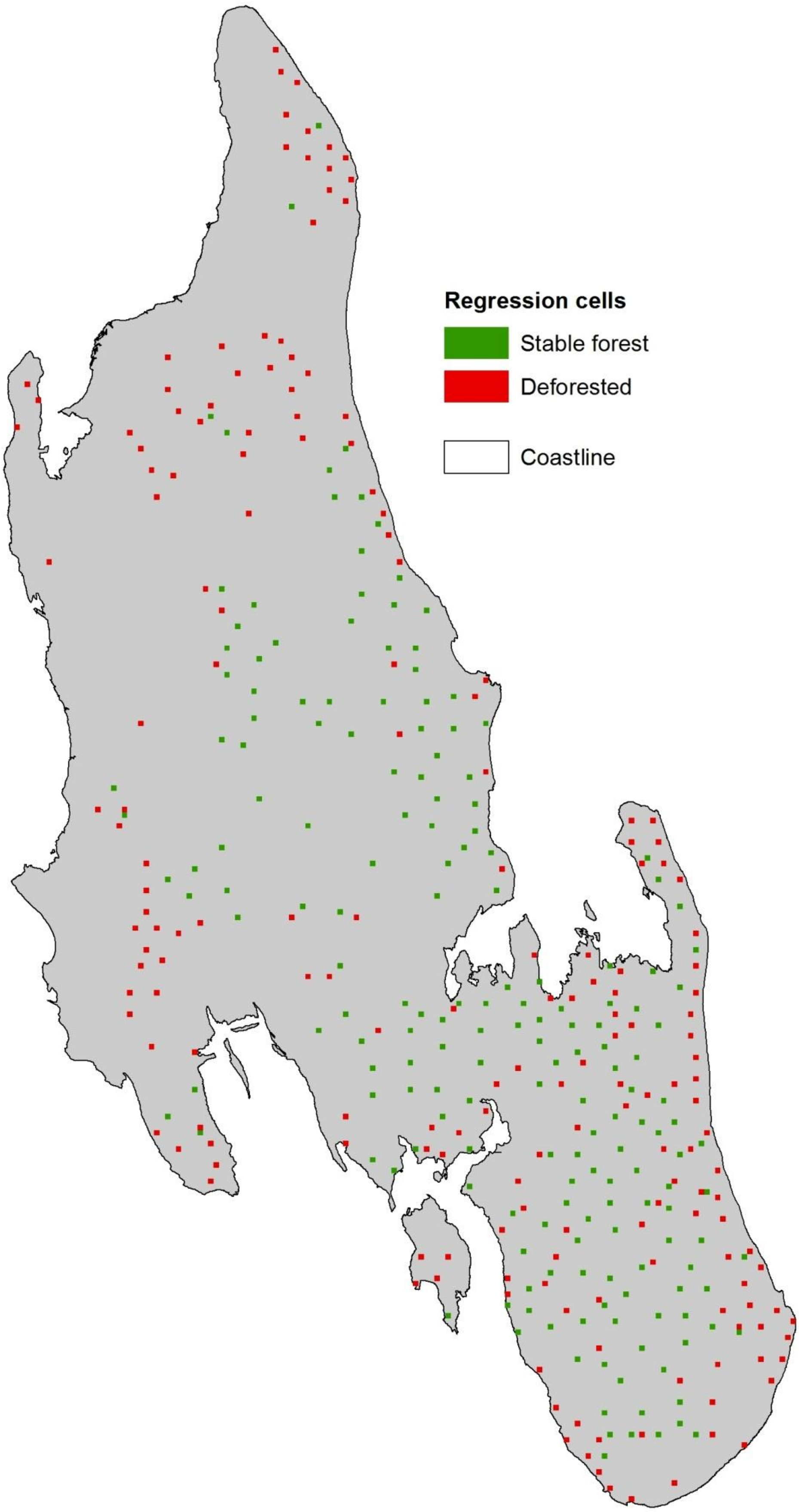
1996 class	2009 class	Gradual class	Abrupt class	Area (km^2)	Area (%)
B	B	Stable non-vegetated	Stable non-vegetated	32,57	2,17
B	U	Stable non-vegetated	Stable non-vegetated	26,36	1,76
B	L-LV	Revegetation	Revegetation	15,66	1,04
B	S-O B	Revegetation	Revegetation	25,86	1,72
B	S-O G	Revegetation	Revegetation	11,88	0,79
B	W	Reforestation	Reforestation	2,16	0,14
B	C F/S	Reforestation	Reforestation	4,42	0,30
U	B	Stable non-vegetated	Stable non-vegetated	1,78	0,12
U	U	Stable non-vegetated	Stable non-vegetated	11,22	0,75
U	L-LV	Revegetation	Revegetation	0,84	0,06
U	S-O B	Revegetation	Revegetation	1,44	0,10
U	S-O G	Revegetation	Revegetation	0,82	0,05
U	W	Reforestation	Reforestation	0,25	0,02
U	C F/S	Reforestation	Reforestation	1,06	0,07
L-LV	B	Devegetation	Devegetation	18,33	1,22
L-LV	U	Devegetation	Devegetation	13,61	0,91
L-LV	L-LV	Stable vegetation	Stable vegetation	37,82	2,52
L-LV	S-O B	Stable vegetation	Stable vegetation	44,44	2,96
L-LV	S-O G	Improved vegetation	Stable vegetation	42,97	2,87
L-LV	W	Reforestation	Reforestation	4,62	0,31
L-LV	C F/S	Reforestation	Reforestation	14,23	0,95
S-O B	B	Devegetation	Devegetation	21,69	1,45
S-O B	U	Devegetation	Devegetation	16,95	1,13
S-O B	L-LV	Stable vegetation	Stable vegetation	29,12	1,94
S-O B	S-O B	Stable vegetation	Stable vegetation	54,58	3,64
S-O B	S-O G	Improved vegetation	Stable vegetation	35,44	2,36
S-O B	W	Reforestation	Reforestation	14,07	0,94
S-O B	C F/S	Reforestation	Reforestation	22,92	1,53
S-O G	B	Devegetation	Devegetation	9,55	0,64
S-O G	U	Devegetation	Devegetation	10,36	0,69
S-O G	L-LV	Vegetation degradation	Stable vegetation	49,35	3,29
S-O G	S-O B	Vegetation degradation	Stable vegetation	54,54	3,64
S-O G	S-O G	Stable vegetation	Stable vegetation	128,65	8,58
S-O G	W	Improved forest	Reforestation	16,55	1,10
S-O G	C F/S	Improved forest	Reforestation	55,23	3,68
W	B	Deforestation	Deforestation	2,84	0,19
W	U	Deforestation	Deforestation	3,29	0,22
W	L-LV	Deforestation	Deforestation	6,40	0,43
W	S-O B	Deforestation	Deforestation	23,27	1,55
W	S-O G	Forest degradation	Deforestation	11,13	0,74
W	W	Stable forest	Stable forest	38,50	2,57
W	C F/S	Improved forest	Improved forest	15,19	1,01
C F/S	B	Deforestation	Deforestation	6,69	0,45
C F/S	U	Deforestation	Deforestation	9,14	0,61
C F/S	L-LV	Deforestation	Deforestation	28,68	1,91
C F/S	S-O B	Deforestation	Deforestation	53,76	3,59
C F/S	S-O G	Forest degradation	Deforestation	86,51	5,77
C F/S	W	Forest degradation	Forest degradation	48,12	3,21
C F/S	C F/S	Stable forest	Stable forest	334,69	22,32

B = Bare, U = Urban, L-LV = Low-lyinh vegetation, S-O B = Semi-open scrubs on barren, S-O G = Semi-open scrubs on grass, W = Woodland, C F/S = Closed forest/scrub.

APPENDIX 3. Land cover changes in Unguja between 1996 and 2009:



APPENDIX 4. The cells used for regression analyses:



APPENDIX 5. Bivariate Pearsons correlation with two-tailed test of statistical significance for Coral rag region

		Dist_mang	1996_ndvi	Ker_buildi	Dist_Stown	GPF_y_n	GFF_y_n	Soil_type	Dist_ms_rd	Elev_mean
Dist_mang	Pearson Correlation	1	,068*	-,039	-,151**	,073*	,100**	,003	,204**	,278**
	Sig. (2-tailed)		,034	,219	,000	,023	,002	,918	,000	,000
1996_ndvi	Pearson Correlation	,068*	1	-,110**	-,256**	,190**	-,016	,040	,244**	,114**
	Sig. (2-tailed)	,034		,001	,000	,000	,624	,209	,000	,000
Ker_buildi	Pearson Correlation	-,039	-,110**	1	,012	-,122**	,038	,017	-,219**	,467**
	Sig. (2-tailed)	,219	,001		,720	,000	,239	,604	,000	,000
Dist_Stown	Pearson Correlation	-,151**	-,256**	,012	1	-,479**	-,066*	,007	,146**	-,068*
	Sig. (2-tailed)	,000	,000	,720		,000	,039	,835	,000	,034
GPF_y_n	Pearson Correlation	,073*	,190**	-,122**	-,479**	1	-,011	,005	-,047	,035
	Sig. (2-tailed)	,023	,000	,000	,000		,732	,879	,141	,279
GFF_y_n	Pearson Correlation	,100**	-,016	,038	-,066*	-,011	1	,003	-,017	,048
	Sig. (2-tailed)	,002	,624	,239	,039	,732		,915	,593	,134
Soil_type	Pearson Correlation	,003	,040	,017	,007	,005	,003	1	,104**	-,026
	Sig. (2-tailed)	,918	,209	,604	,835	,879	,915		,001	,424
Dist_ms_rd	Pearson Correlation	,204**	,244**	-,219**	,146**	-,047	-,017	,104**	1	,219**
	Sig. (2-tailed)	,000	,000	,000	,000	,141	,593	,001		,000
Elev_mean	Pearson Correlation	,278**	,114**	,467**	-,068*	,035	,048	-,026	,219**	1
	Sig. (2-tailed)	,000	,000	,000	,034	,279	,134	,424	,000	

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

APPENDIX 6. Bivariate Pearsons correlation with two-tailed test of statistical significance for Deep soil region

		Dist_mang	1996_ndvi	Ker_buildi	GPF_y_n	GFF_y_n	Dist_Stown	Soil_type	Dist_ms_rd	Elev_mean
Dist_mang	Pearson Correlation	1	-,326**	-,107*	-,181**	-,065	,093	-,520**	-,137**	,534**
	Sig. (2-tailed)		,000	,040	,000	,213	,076	,000	,009	,000
1996_ndvi	Pearson Correlation	-,326**	1	-,084	,021	-,180**	-,192**	,465**	-,141**	-,280**
	Sig. (2-tailed)	,000		,108	,692	,001	,000	,000	,007	,000
Ker_buildi	Pearson Correlation	-,107*	-,084	1	,257**	-,129*	-,627**	-,283**	,098	-,184**
	Sig. (2-tailed)	,040	,108		,000	,014	,000	,000	,062	,000
GPF_y_n	Pearson Correlation	-,181**	,021	,257**	1	-,069	-,273**	-,085	,153**	,270**
	Sig. (2-tailed)	,000	,692	,000		,184	,000	,103	,003	,000
GFF_y_n	Pearson Correlation	-,065	-,180**	-,129*	-,069	1	,510**	,083	,431**	,187**
	Sig. (2-tailed)	,213	,001	,014	,184		,000	,110	,000	,000
Dist_Stown	Pearson Correlation	,093	-,192**	-,627**	-,273**	,510**	1	,013	,249**	,382**
	Sig. (2-tailed)	,076	,000	,000	,000	,000		,803	,000	,000
Soil_type	Pearson Correlation	-,520**	,465**	-,283**	-,085	,083	,013	1	,021	-,447**
	Sig. (2-tailed)	,000	,000	,000	,103	,110	,803		,686	,000
Dist_ms_rd	Pearson Correlation	-,137**	-,141**	,098	,153**	,431**	,249**	,021	1	,102*
	Sig. (2-tailed)	,009	,007	,062	,003	,000	,000	,686		,050
Elev_mean	Pearson Correlation	,534**	-,280**	-,184**	,270**	,187**	,382**	-,447**	,102*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,050	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).