**Introduction**

Southeast Asia (SEA) is characterised by complex biogeography and extensive tropical forest cover resulting in exceptional biological diversity, yet it has one of the highest rates of deforestation in the world (Hughes, 2017). Deforestation rates in SEA are comparable only with those of Latin America (Estoque et al., 2019), and the resulting habitat loss is arguably more damaging to biodiversity (Sodhi et al., 2010, 2004). The loss of SEA’s forests has potentially severe consequences for climate change (Ceddia et al., 2015), ecosystem-based adaptation (Estoque et al., 2019), local people (Culas, 2007; Frewer and Chan, 2014; Gaughan et al., 2009; Poffenberger, 2006), and biodiversity (Chapman et al., 2018; Hearn et al., 2018). The drivers of tropical deforestation vary both by location and by scale, ranging from broader drivers such as population pressure and weak institutions (Geist and Lambin, 2002), to proximate causes at a local level such as the expansion of cash crops, agriculture, and other food production (Estoque et al., 2019; Imai et al., 2018; Stibig et al., 2014; Wilcove et al., 2013; Zeng et al., 2018), the associated development of roads and infrastructure that facilitate such expansion (Hughes, 2018), and civil unrest and war (Kaimowitz and Fauné, 2003; Price, 2020).

Deforestation is accentuated in low income countries where poverty, debt, and overpopulation are high, and thus so are the demands for economic growth via agricultural commodities and other natural resource-based products (Culas, 2012). Research has shown that this trend is reversed in high income countries where economic activity shifts to become more service-driven, and demand for environmental services increases, resulting in afforestation (Culas, 2012). The change from deforestation (low income) to afforestation (high income) is termed forest transition (Mather, 1992). There is also evidence that remaining forest area is an effective predictor of deforestation rates, with countries undergoing forest transition when remaining forest cover across the country is low (Lambin and Meyfroidt, 2010). Within SEA, countries such as Vietnam have undergone forest transitions and reduced deforestation rates dramatically over the last 20 years (Meyfroidt and Lambin, 2008). Yet often this progress is at the expense of poorer, less developed countries with weaker institutions and policy frameworks, such as Cambodia, where activities including logging and forest clearance leak across borders (Meyfroidt and Lambin, 2009).

There are various broad pathways which can result in a land use transition within a country (Lambin and Meyfroidt, 2010). Despite the terminology, these pathways are not linear or deterministic; they are driven by complex endogenous and exogenous drivers and feedback loops between economic, political, social, and environmental factors operating at a variety of scales (Lambin and Meyfroidt, 2010). To facilitate the development of sustainable land use policies that can guide countries towards more desirable forest transition pathways, and to support targeted interventions that are effective at reducing forest loss, it is critical to identify and understand macroeconomic and socioeconomic drivers and predictors of forest cover and loss at different scales. Describing relationships between forest cover and macroeconomic and socioeconomic variables, rather than simply biophysical variables, is particularly important in developing countries that are undergoing social transition, as these are the conditions under which deforestation is often accelerated (Imai et al., 2018). Successfully isolating the signals of these relationships is however, challenging, due to the complexity of social-ecological systems, the non-linear feedback loops, and the heterogeneity in system dynamics at different scales.

***Modelling approaches***

Modelling the processes that drive land use change (LUC), including forest loss, is challenging due to the complexity of the systems in which it occurs (Basse et al., 2014). Advanced modelling frameworks have been developed that are allowing greater understanding of the processes underlying LUC, and subsequently more accurate predictions into the future (Basse et al., 2014; Bonilla-Bedoya et al., 2018). However, the modelling approach taken depends on several factors including 1) the research question, e.g., whether you want to model *patterns* of LUC or the *processes* of LUC, 2) the scale of the analysis (Brown et al., 2013), e.g., understanding human decision-making that drives land conversion at the local level, or understanding how global macroeconomics affect agricultural expansion at a regional or national level, 3) expertise and interests, which govern both the research questions and the statistical approaches that are likely to be used, and 4) data availability, which although has seen dramatic improvements in recent decades thanks to freely available, high resolution remote sensing imagery and long-term global data sets on socioeconomics, can still limit analyses, particularly when there are scale mismatches between data (Ceddia, 2019).

Methodological approaches fall broadly into two groups governed largely by the aims of the study. First, modelling the spatial processes of LUC is a common goal, as this allows researchers to use patterns of past LUC to predict which areas are at higher risk of land conversion in the future, with the potential to explore a number of plausible future scenarios (Basse et al., 2014). There are several spatially explicit cell-based modelling frameworks that can achieve these aims, including maximum entropy (Bonilla-Bedoya et al., 2018; de Souza and De Marco, 2014), and cellular automata (Stevens and Dragićević, 2007; Yang et al., 2012), which rely on discrete spatial units that have associated variable values and tend to be spatially correlated. Models are trained on past LUC, thus developing rules which govern the likelihood of a given cell being converted under various future conditions. These rule-based approaches can be improved with advanced machine learning techniques such as artificial neural networks (Basse et al., 2014). More traditional statistical models have also been used in a cell-based framework, whereby models such as logistic regression are used to predict the likelihood of a cell being converted (Aguiar et al., 2016; Aspinall, 2004). The above methods require spatially explicit data at the scale of the cell, and so are unable to model broader relationships and drivers that operate at a different level within the system. These approaches rely on geographical and biophysical predictor variables that have fixed values in space, precluding the investigation of national or regional drivers. Socioeconomic drivers can be modelled, provided they can be represented in discrete space and are at the correct scale (de Souza and De Marco, 2014; Estoque et al., 2019).

Second, researchers may want to model the relationships between LUC and trends in predictor variables over time rather than space. These approaches are generally less deterministic than the spatial process modelling above, and are often at much larger scales (e.g., Bhattarai and Hammig, 2004; Ceddia, 2019; Ewers, 2006). These analyses are often targeting broader economic, socioeconomic, cultural, political, and institutional drivers of LUC, which are less amenable to spatial sampling. Generalised linear mixed models (GLMMs, also known as multilevel or hierarchical models) are often employed in such analyses, as GLMMs can account for temporal autocorrelation and hierarchical data structures (Zuur et al., 2009). Studies have used these, and other regression-type models, to investigate the relationships between LUC and national income and forest policies (Bhattarai and Hammig, 2004), income, land, and wealth inequalities (Ceddia, 2019), indigenous land tenure (Ceddia et al., 2015), macroeconomics and economic development (Culas, 2007; Ewers, 2006), and urban socioeconomics (Gong et al., 2013). Studies that use GLMMs almost exclusively use data from multiple countries, taking advantage of the ability of these models to harness large longitudinal data sets with few “subjects” without succumbing to pseudoreplication (Gelman and Hill, 2006, see Ewers, 2006 for an example of where pseudoreplication may be an issue). Another advantage of GLMMs is the ability to quantify between-group variance, which not only offers crucial insight about the differences between groups (e.g., countries) from which inference can be drawn (Zuur et al., 2009), but can also highlight potential problems with ‘global’ predictions (i.e., predictions that are made with all random effect terms set at their mean). Yet very few studies that use these models for LUC report any values for variance associated with the random (group-level) effects. For example, Bhattari and Hammig (2004) use data from 63 countries to produce a single effect for GDP per capita on deforestation, yet do not report any value for country-level variance. The effect size is relatively small, and therefore if there was large between-country variance then the country-level effects could be vastly different, rendering the single global effect misleading. Furthermore, some studies fail to even specify the statistical or modelling approaches used to arrive at their conclusions (e.g., Gao and Liu, 2012).

Modelling complex, non-linear systems that interact and feedback at various temporal and spatial scales is difficult (Dawson et al., 2010), and conservation challenges can arise when there are mismatches in the interactions and scales between social and ecological systems (Beever et al., 2019). It is therefore important for researchers to both accept and report high levels of uncertainty and variation within models that attempt to simplify complex systems. For example, without appropriate estimates of variance at the different levels within GLMMs, it is difficult to reliably assess model coefficients. This is particularly true when hierarchical model levels represent different spatial scales (e.g., countries within regions, cities within states).

***Cambodia***

Between 1975 and 1992 Cambodia suffered enormous civil unrest, war, and foreign occupation, which resulted in almost complete economic collapse. Yet by 2006 Cambodia’s economy was one of the fastest growing economies in the world and represents a good example of a country’s ability to move from post-conflict status to full integration within a dynamic regional economy (Hughes and Un, 2011). This remarkable economic recovery has, however, come at a cost for the country’s natural environment (Davis et al., 2015).

In the decades following the 1991 Paris Peace accords, which officially ended civil war and foreign occupation in Cambodia, political stability and rapid economic growth resulted in an economic transformation (Hughes and Un, 2011). Between 1990 and 2006 the economy transitioned from 85% subsistence agriculture, a small garments sector, and generally low economic productivity, to 55% of the population involved in subsistence agriculture and a private sector that was growing exponentially, thanks to investments in tourism, manufacturing, and mining (Eliste and Zorya, 2015; Hughes and Un, 2011). Between 2000 and 2006 the economy grew by an average of 8.7% - one of the highest rates in the world - driven primarily by manufacturing, (especially garment manufacturing), construction, services, and tourism. These industries were geographically limited to the two major cities: Phnom Penh, Siem Reap, and their surrounds. Outside of the large cities, the rural economy was also undergoing significant change. The adoption of the 2001 Land and Forestry Laws saw a major drive towards industrial-scale agriculture in the form of private and state-owned land concessions targeting commercial crops such as rubber and sugar (Eliste and Zorya, 2015). These concessions drove most of the growth within the agricultural sector, and despite having positive economic impacts, have received heavy criticism for violations of local land rights and illegal deforestation within protected areas (Global Witness, 2013; Vrieze and Kuch, 2012; Watson et al., 2014).

The rapid economic development around the large urban centres has increased the gap in development and socioeconomic status between the urban and rural populations. Despite the national economic success since the 1990s, Cambodia is still one of the least developed countries in the world, with more than 33% of the population either in severe poverty or vulnerable to severe poverty (UNDP, 2020). In one study of a rural province in Cambodia, it was found that 70% of the population were either low-skilled non-permanent wage employees, permanent farmers, or were involved in resources extraction (Nguyen et al., 2015). The authors further found that all households, even those with high skilled or permanent wage employees, engaged in personal agriculture or livestock rearing, and were in some part dependent on environmental resources (Nguyen et al., 2015). Rural poverty and reliance on natural resources have been exacerbated by decades of insecure land tenure, followed by several ambitious, yet poorly governed, land tenure policies such as Directive 01 (Grimsditch and Schoenberger, 2015; Milne, 2013). These land tenure policies resulted in widespread land speculation and rapid deforestation across the country, including within protected areas (see Thesis appendix).

At the start of the century 41.9% of Cambodia’s land area was forested, and by 2012 the total forested area had been reduced by 19.8%, equating to over 1.3 million hectares (Davis et al 2015). Only 25 other countries lost more forest than Cambodia between 2000 – 2012 (Hansen et al 2013). Such high rates of deforestation, in the context of Cambodia’s economic status and the rural population’s reliance on natural resources (Nguyen et al., 2015), suggests that the country is on the increasing trajectory of the Environmental Kuznet’s Curve, whereby national economic development is improved at the expense of forest cover (Stern, 2004). This theory is supported by evidence of increased forest loss within agro-industrial land concessions, highlighting these economic concessions as key drivers of deforestation (Davis et al., 2015). These Economic Land Concessions (ELCs) have been an important, and controversial, tool in the economic recovery of Cambodia. Nevertheless, the importance of natural resources for the country’s large rural population, and in particular the reliance on forest resources by the country’s indigenous populations (Nguyen et al., 2015), means that reducing forest loss and directing the country towards a forest transition is important (Rudel et al., 2005). Therefore, understanding the relationships between macroeconomics, socioeconomics, and forest cover at different scales across the country will be critical to inform sustainable economic policies and effective forest conservation programmes.

***Purpose of the study***

In this study we investigate the relationships between forest cover and loss, and macroeconomic and socioeconomic variables in Cambodia. Specifically, we: 1) model the relationships between three sets of macroeconomic predictors and forest loss at the national scale, 2) model the relationships between three sets of macroeconomic predictors and the creation of new ELCs at the national scale, 3) model the relationships between forest cover and eight sets of socioeconomic predictors, at two different sub-national scales, and 4) conduct a cluster analysis to group provinces by socioeconomics in order to further understand the socioeconomic typologies within the county, and their relationships to forest cover.

**Methods**

***Study area***

Map

Description automatically generatedThis study investigated macroeconomic and socioeconomic predictors of forest cover and loss in Cambodia at two different spatial resolutions over two time periods. The macroeconomic analysis was at the national scale between 1993 – 2015, and the socioeconomic analysis was at the scale of a) the commune, and b) the province between 2007 – 2012. Cambodia is in mainland SEA and is bordered by Laos (NE), Thailand (NW), Vietnam (E), and the Gulf of Thailand (SW) (Figure 1). The country has a surface area of 176,520 km2 (UNCTAD, 2020) and is located at latitudes 10-14° north of the equator and thus has a tropical monsoon climate (McSweeney et al. 2010).

**Figure 1. Map of Southeast Asia with Cambodia highlighted in red.**

***Variable selection***

The response variables for the macroeconomic analysis were 1) change in forest cover (forest loss) from time *t* to time *t+1* and 2) the number of new economic land concession (ELC) allocations in year *t*. The response variable for the socioeconomic analysis was forest cover area. Both forest cover response variables were produced from the same data source (see “Data sources” below). Macroeconomic, socioeconomic, and control variables for both sets of analysis were selected based on a combination of previous studies, data availability, and the authors’ knowledge of Cambodia.

Macroeconomic variables were selected to create three sets of predictors, each targeting a different driver: economic development (n=8), commodity prices (external market forces, n=8), and producer prices (internal market forces, n=5) (Table S1, Nelson et al. 2006; Ewers 2006; Gong et al. 2013; Kuang et al. 2016; Fan & Ding 2016; Bonilla-Bedoya et al. 2018). Each predictor was hypothesised to be a driver of forest loss (Table S2). Human population density was included as a control variable for the economic set and total forest remaining was included as control variable across all sets, as both were expected to influence forest loss. Both per capita Gross Domestic Product (GDP) and amount of forest remaining were included to reflect the economic development path and the forest scarcity path respectively (Lambin and Meyfroidt, 2010; Rudel et al., 2005). After pre-analysis checks for errors and correlation, the resulting variable set contained 20 variables (Table 1).

Socioeconomic variables were selected to create 8 variable sets reflecting different aspects of socioeconomic status and development, each of which was hypothesised to be either a driver or predictor of forest cover (Table S3, Luck et al. 2009; Ty et al. 2012; Kristensen et al. 2016; Bonilla-Bedoya et al. 2018). The variable sets were population demographics (n=8), education (n=4), employment (n=5), economic security (n=2), access to services (n=4), social justice (n=2), migration (n=2), and control variables (n=6). Control variables were included to account for the effects of environmental and other human factors including economic land concessions (Abdullah and Nakagoshi, 2007; Davis et al., 2015), protected areas (Bonilla-Bedoya et al., 2018), elevation (Ty et al., 2012), and distance to human infrastructure (Ty et al., 2012). A habitat control variable was excluded because the response variable (forest cover) was extracted from a land cover layer and represented a specific type of habitat, resulting in non-independence between the response and habitat.

***Data sources***

National macroeconomic variables were acquired from publicly available sources (Table 1) for the period 1993 – 2015. Fine-scale socioeconomic variables were extracted from the Cambodian Commune Database for the years 2007 – 2012 (Table 2) which are available from Open Development Cambodia ([www.opendevelopmentcambodia.net](http://www.opendevelopmentcambodia.net)). Data on economic land concessions, protected areas, and elevation (digital elevation model), and shapefiles for the country, provinces, and communes were provided by the Wildlife Conservation Society. Forest cover layers were taken from the publicly available European Space Agency Climate Change Initiative (ESACCI) satellite data for the years 1993 – 2015.

**Table 1. Variables selected for the macroeconomic analysis. Variables range from 1993 – 2015.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Predictor variable** | **Units** | **Resolution** | **Source** | **Details** |
| *Economy* |  |  |  |  |
| GDP per capita | Billions USD | National | World Bank | Constant 2010 rates |
| GDP growth | % | National | World Bank | Annual percentage growth rate of GDP at market prices based on constant local currency |
| Foreign Direct Investment | Millions USD | National | UNCTAD | Inward and outward flows and stock |
| Agricultural sector proportion of GDP | % | National | CNIS | Proportion of national GDP |
| Development flows to agriculture | Millions USD | National | FAO | Donor development investment flows, other official flows, and private donor flows at constant 2016 prices to all agriculture and forestry sub-sectors |
| Development flows to environment | Millions USD | National | FAO | Donor development investment flows, other official flows, and private donor flows at constant 2016 prices to general environment protection |
| *Commodity prices* |  |  |  |  |
| Crop Production | Index | National | FAO | Relative level of the aggregate volume of agricultural production for each year in comparison with the base period 2004-2006 |
| Non-food agricultural production | Index | National | FAO | Relative level of the aggregate volume of non-food agricultural production for each year in comparison with the base period 2004-2006 |
| Forestry production | m3 | National | FAO | Total production values for industrial roundwood, non-coniferous tropical wood, other industrial roundwood, sawlogs and veneer logs (coniferous and non-coniferous), and sawnwood (coniferous and non-coniferous |
| Price of rice | USD/ton | Global | World Bank | Median annual global market price of rice |
| Price of corn | USD/ton | Global | World Bank | Annual global market price of corn |
| Price of rubber | USD/ton | Regional | RASCE | Monthly regional market value of rubber on the Singapore Exchange |
| Price of sugar | USD/ton | Global | World Bank | Annual global market price of sugar |
| *Producer prices* |  |  |  |  |
| Producer price of Rice | USD/ton | National | FAO | Farmgate prices for Cambodian producers |
| Producer price of rubber | USD/ton | National | FAO | Farmgate prices for Cambodian producers |
| Producer price of cassava | USD/ton | National | FAO | Farmgate prices for Cambodian producers |
| Producer price of corn | USD/ton | National | FAO | Farmgate prices for Cambodian producers |
| Producer price of sugar | USD/ton | National | FAO | Farmgate prices for Cambodian producers |
| *Control* |  |  |  |  |
| Population density | pax/km2 | National | FAO | People per km2 |
| Forest remaining | Km2 | National | ESACCI | Raw value of forest remaining |

**Table 2. Variables selected for the socioeconomic analysis. Variables range from 2007 – 2012.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Set** | **Variable** | **Transformation for analysis** | **Province-level class** | **Details** |
| Demographics | Total population |  | NA | Includes women, men, and children of all ages |
|  | Population density |  |  |  |
|  | Number indigenous | Proportion of total population | Categorical | Total number of people who are indigenous/ethnic minority |
| Education | Males aged 6 – 24 in school | Proportion of total number of males aged 6 - 24 |  | Number of males aged 6 - 24 in full time education |
| Employment | Number of adults employed in primary sector | Proportion of total adult population | Categorical | The primary sector includes agriculture (rice and other crop farming), fishing, livestock farming, forestry, and non-timber forest product collection (Kenessey 1987) |
|  | Number of adults employed in secondary sector | Proportion of total adult population | Categorical | The secondary sector includes wood-based production (e.g. furniture), metal- and glass-based production, foodstuff production, plastic- and rubber-based production, textiles production (Kenessey 1987) |
| Economic security | Number of families with <1ha rice land (including no rice land) | Proportion of total number of families | Categorical |  |
|  | Number of families who keep pigs | Proportion of total number of families | Categorical |  |
| Access to services | Distance to nearest school |  | Categorical | Median distance from any village in the commune to the nearest school (primary or secondary) |
|  | Number of families with access to waste collection | Proportion of total number of families |  |  |
|  | Distance to the Commune Office |  |  | Median distance from any village in the commune to the Commune Office (government administration office) |
| Social justice | Number of criminal cases | Criminal cases per capita | Categorical | Includes murder, theft, and other criminal cases |
|  |  |  |  |  |
|  | Number of land conflict cases |  | Categorical | In the previous 12 months |
| Migration | Number of in-migrants |  | Categorical | Migration into the commune |
|  | Number of out-migrants |  | Categorical | Migration out of the commune |
| Control | Mean elevation (masl) |  | Categorical | Mean elevation for the commune |
|  | Distance to international border (km) |  | Categorical | Distance from the centre of the commune to the nearest international border |
|  | Distance to Provincial Capital (km) |  | Categorical | Distance from the centre of the commune to the centre of the provincial capital (town or city) |
|  | Presence of economic land concessions |  |  | Binary. 1 = part or all of an economic land concession falls within the boundary of the commune, 0 = no economic land concession falls within the commune boundary |
|  | Presence of protected area |  |  | Binary. 1 = part or all of an protected area falls within the boundary of the commune, 0 = no protected area falls within the commune boundary. "Protected area" includes Wildlife Sanctuary, National Park, Protected Landscapes, Multiple-use areas, RAMSAR sites |
|  | Protected area category |  |  | None = no protected area falls within commune, MULTI = more than one category of protected area falls within commune, WS = wildlife sanctuary, NP = national park, PL = protected landscape, MUA = multiple-use area, RMS = RAMSAR |

***Data processing***

The forest cover variable (response) for both analyses was extracted from the ESACCI product by totalling the number of pixels (1 Pixel = 0.09km2) in each year classified as bands 50, 60, 61, 62, 70, 71, 72, 80, 81, 82, 90, and 100 (Table S4). For the macroeconomic analysis, the total forest cover for the whole country was used, and for the socioeconomic analysis the forest cover layer was further stratified into forest cover per commune and forest cover per province. Forest cover data processing was done in QGIS (QGIS Geographic Information System v3.16). For both analyses, predictor variables were checked for collinearity, and if two variables in the same set had a correlation coefficient of >0.6 then generally one was removed (Supporting Information).

*Macroeconomic analysis*

Forest cover was converted to change in forest cover using *forest covert+1 − forest covert*, where *t* represents year *t.* There were no periods of forest gain during the study period, and so the response can be considered as rate of forest loss. All predictors were converted from raw values to change in values using *Xt+1 – Xt,* where *t* represents year *t* (Barrett et al., 2006). The variable *forest remaining* was left as raw values (km2). Cambodia’s first general election and subsequent adoption of a free market economy occurred in 1993, resulting in unreliable GDP-related values for 1993 (Chhair and Ung, 2013) and subsequent change values in 1994, and so these were removed. To simplify interpretation, predictor variables were not centred or scaled prior to change calculations or modelling.

*Socioeconomic analysis*

Data from the Commune Database were at the resolution of individual village, and so the selected variables (Table 2) were aggregated to the commune and province level after error checking and cleaning (Supporting Information). This resulted in between 1,317 and 1,512 communes, and 23 Provinces (excluding Phnom Penh). The number of communes changed between years due to administrative changes. Some variables were converted from raw values to proportional data to account for large differences in commune and province size and human population (Table 2). Data were checked for errors in R (Supporting Information, R Core Team, version 4.0).

***Modelling***

*Macroeconomic models*

This analysis aimed to model the relationships between changes in macroeconomic predictors and 1) the change in forest cover at a national level and 2) the allocation of new ELCs, for the time period 1993 – 2015. Models were run for both response variables with each of the three variable sets: economic development, commodity prices, and producer prices. To account for the effect of time, a linear model of the response as a function of time (year) was run and the model residuals were extracted and used as a control predictor in all subsequent models. The amount of forest remaining (km2) was also included as a control variable in all models. Modelling was done using Generalised linear models (GLM) and followed an information theoretic approach (Burnham and Anderson, 2007). For the models with rate of forest loss as the response both gaussian and gamma distributions were tested, and for the models with ELC allocation a Poisson distribution was used. Resulting models were compared using Akaike’s Information Criterion (AIC). Final rate of forest loss models used gaussian distributions. All predictors in each model set had been selected because of a priori hypotheses (Table S2), and so within each set all combinations of possible models were run and compared using AIC. Models with ∆AIC < 6 were considered to have sufficient support and retained in the final model set. Model averaging was implemented for the final model set, resulting in model-averaged coefficients for all model terms (Burnham and Anderson, 2007). Models were run and averaged using the MuMIn package in R (Version 1.43.17, Bartoń 2020). This modelling procedure was repeated for a one-year time lag and two-year time lag as follows:

No time lag:

Where is the response at time , and is predictor variable at time.

One year time lag:

Where is the response at time , and is predictor variable at time.

Two year time lag:

Where is the response at time , and is predictor variable at time.

*Socioeconomic models*

This analysis aimed to model the relationships between forest cover and socioeconomic variables within communes between 2007 – 2012. The results of initial commune-level modelling prompted further aggregation of the data to the province-level and models were built to investigate the relationships between forest cover and socioeconomic variables within provinces for the same time period.

*Commune-level models*

Generalised linear mixed models (GLMM) with Poisson errors were used with commune nested within province as random intercept terms to account for repeat measurements and the hierarchical data structure, and year as a random slope term to account for temporal autocorrelation (Zuur et al., 2009). The natural logarithm of commune area (km2) was used as an offset term in all models to account for large variation in commune size. Due to the large number of available predictor variables, maximal within-set models were run first for each of the 8 variable sets (Table S8), and variables with very weak, or no effect were dropped. Simplified models were compared with maximal models using likelihood ratio tests and analysis of variance tests. If a variable set had only one variable, this was automatically taken forward. Because assessment of term significance in GLMMs is complex, predictions and plots were made for all terms before being dropped to ensure noteworthy effects were not being missed. This process resulted in a final set of 13 variables which were used to create a candidate set of 10 models (Table S19). Following an information theoretic approach (Burnham and Anderson, 2007) models were compared via AIC to select the top model or models. The resulting final model fit was assessed via diagnostic plots (residuals versus fitted, quantile-quantile of random effects, Supporting Information, Harrison et al (2018)). Marginal (fixed effects only) and conditional (fixed and random effects) pseudo-R2 values were calculated based on Nakagawa & Schielzeth (2017) using the R package ‘MuMIn’ (Bartoń, 2020). To investigate the variation in effects between provinces, predictions were made for each variable within each commune and the 50% quantile from all commune-level predictions within each province was extracted as the provincial mean prediction.

*Province-level models*

The same GLMM model formulation was used for the province-level models except that commune was removed from the random effects structure. Based on provincial-level histograms of predictor variables, 14 predictors were converted to categorical variables by splitting the data by the mean, resulting in “high” and “low” values (Table 2). Following an information theoretic approach, a candidate set of models was created (Table Sx) and model comparison was done using AIC.

***Cluster analysis***

Agglomerative clustering was conducted to create a typology for provinces based on the socioeconomic variables used in the analysis above. Several agglomerative clustering approaches were assessed. These were single linkage, complete linkage, unweighted pair-group using arithmetic averages (UPGMA), unweighted pair-group using centroids (UPGMC), Ward’s minimum variance, and flexible clustering. The methods were compared using cophenetic correlation and Gower distance metrics, and the appropriate number of clusters (k) was selected using the matrix correlation statistics (Borcard et al., 2018). The capital city of Phnom Penh, which is technically a province in itself, was removed prior to clustering because it has extreme values for many of the variables and is thus an outlier that affects the clustering.

**Results**

***Macroeconomic analysis***

*Rate of forest loss response*

Models revealed that there were no strong effects of the macroeconomic predictors on forest loss between 1993 and 2015 (Figures S2 – S4). For each predictor set there were between 5 and 28 models in the top model set and final coefficients were calculated using full averages (Tables S9 – S17, Burnham and Anderson, 2007). The largest effects were from two of the control variables (population density, time). The largest effect overall was for population density with a one-year time lag (full averaged coefficient = -632.9, SE = 64.8, Table S10). The largest effect excluding control variables was for agricultural proportion of GDP with a one-year time lag (full averaged coefficient = -14.9, SE = 7.9) suggesting that there is a small reduction in the rate of forest loss as the contribution of agriculture to national GDP increases, although this effect is very weak (Figure S2, Table S10) and is not considered definitive.

*New economic land concession response*

There were 287 new ELCs allocated within the study period, with the majority (51%) being designated for rubber production (Table S18). The most valuable crop in terms of commodity price during the study period was rubber, with a mean market price of $1743/ton, followed by rice ($348/ton) and sugar ($282/ton, Table S18). The most valuable crop in terms of producer (farmgate) prices was sugar with a mean price over the study period of $2115/ton, followed by rubber ($317/ton) and corn ($197/ton, Table S18). A greater number of effects were revealed in the macroeconomic analysis with new economic land concession allocation as the response. The largest effect was for the economic control variable population density, where there were very strong negative effects across all time lags (rate ratios for one-year lag = 0.012, two-year lag = 0.002, three-year lag = 0.0005, Table 3), indicating that new ELCs do not get allocated in areas of high human population density. The largest overall effect excluding control variables was for changes in agricultural proportion of GDP with no time lag and a one-year time lag (no time lag rate ratio = 1.310, and one-year time lag rate ratio = 1.284, Table 3, Figure 2).

From an economic perspective there were positive relationships between the allocation of new ELCs and increases in the agricultural proportion of GDP and increases in foreign direct investment (one-year time lag rate ratio = 1.004, Table 3, Figure 2). These effects suggest ties between both the development of new industrial-scale concessions and the growth of the agricultural sector, and the injection of foreign wealth into the sector via the purchasing of concessions by international companies. For example, when the agricultural sector’s proportion of national GDP decreases by 3% in a given year relative to the previous year, the number of new ELCs allocated that year is predicted to be approximately 2, whereas when the sector’s proportion of national GDP increases in a given year by 1% relative to the previous year, the number of new ELCs is predicted to be 6. When the amount of foreign investment decreases by approximately $10 million relative to the previous year, the number of new ELCs one year later is predicted to be 3. Conversely, when foreign investment in a given year increases by approximately $300 million relative to the previous year, then one year later the number of new ELCs is predicted to be 10. The one-year time lag of the effect of foreign investment suggests that it takes approximately one year from the time of investment for a company to see the creation of their land concession. There was also a positive relationship between new ELC allocation and increases in development flows to the environment sector (no time lag rate ratio = 1.031). This suggests that in the short-term, investments into the environment sector via development funding (predominantly from international donors) does not reduce the number of new ELC allocations.

There was a negative relationship between new ELC allocation and increases in per capita GDP (one-year time lag rate ratio = 0.985 and two-year time lag rate ratio = 0.974, Table 3, Figure 2). The reduction in ELC allocation as GDP increases, over a period of one and two years, potentially suggests that there is a positive economic effect of ELCs. New concessions inject money into the national economy at various scales, for example at the national level via taxes to the government, and to the local level via employment opportunities and infrastructure development. Thus, as the economy grows, the need for new ELCs diminishes. For example, when GDP per capita in a given year falls by approximately $6 relative to the previous year, the number of new ELCs is predicted to be 8, whereas when the GDP per capita rises in a given year by approximately $60 relative to the previous year, the number of new ELCs predicted is only 3.

The largest effect within the commodity set was for the change in market price of rice in the same year as the response (no time lag) with a rate ratio of 1.009 (Table 3). There were further strong positive relationships between the changes in the market price of rubber (no time lag rate ratio = 1.001), the changes in the non-food production index (one-year time lag rate ratio = 1.007), and changes in the market price of sugar (two-year time lag rate ratio = 1.009). Economic land concessions in Cambodia are predominantly agro-industrial concessions, and therefore the positive relationships between the market price of agricultural commodities and new ELC allocations is not surprising. Rubber and rice are the most valuable market commodities within the variable set, and we can see this reflected in the model; if rubber market prices do not change between years *t* and *t+1* then approximately 4 new ELCs are predicted in year *t+1*, whereas if the price of rubber increases by $1500/ton in year *t*, then approximately 29 new ELCs are predicted in year *t+1*. Similarly, if there is no change in the market price of rice between two given years, then approximately 5 new ELCs are predicted. If the market value increases by $300/ton then in year *t+1* approximately 80 new ELCs are predicted. Interestingly the effect of changes in sugar price were weak when there was no time lag, but the effect was stronger when both a one-year and two-year time lag were introduced (Figure 3).

There were three negative relationships between ELC allocation and commodity variables, all of which were in the same year as the response (no time lag, Figure 3). There were weak negative effects of changes in the market prices of corn (no time lag rate ratio = 0.997) and sugar (no time lag rate ratio = 0.999). Considering the stronger positive effects of sugar price on ELCs after one- and two-year lags, it is unlikely that the very weak negative effect with no time lag is meaningful. The non-food production index had a much stronger negative effect on ELC allocation when there was no time lag (rate ratio = 0.990). The change in direction of the effect of the non-food production index between no time lag and a one-year time lag suggests a complex relationship between the index and ELCs.

The producer price variable set, which reflects the farmgate prices of the commodities, had both positive and negative relationships with ELC allocation (Figure 4, Table 3). The strongest positive relationship was with changes in the producer price of rubber (no time lag rate ratio = 1.035). The effect of positive changes (i.e., net increases) in the price a farmer will get for rubber production can be seen in the predictions of new ELCs (Figure 4). The difference between the number of ELC allocations when the producer price of rubber changes from a decrease of $30/ton (from year *t* to year *t+1*) to no change at all (i.e., the price remains constant) is approximately 3. In contrast, the difference in ELC allocation between no change in price and a positive change of $30/ton is more than 12. This suggests that producers are highly influenced by sale prices of commodities, particularly of high value products such as rubber, and that they will act quickly when there is the potential for financial gain. There were also positive relationships between ELC allocation and changes in the producer price of corn (one-year time lag rate ratio = 1.011) and the producer price of rice (two-year time lag rate ratio = 1.013, Figure 4, Table 3). Corn and rice are less valuable in terms of absolute producer prices than sugar and rubber, and this may be reflected in the time lag that exists between positive changes in the prices and increases in new ELCs.

There were two negative relationships between producer price variables and new ELC allocations (Figure 4). Increases in the producer prices of rice and cassava resulted in fewer predicted ELCs in the same year (no time lag rate ratio = 0.976) and two years later (two-year time lag rate ratio = 0.982), respectively. The difference in the direction of the effect of rice producer prices in year *t* and year *t+2* (Figure 4) suggests that there is a complex relationship between rice production and new ELC allocation. Rice production is the dominant agricultural crop in Cambodia and is the second most valuable commodity in terms of market value (Table S18). Yet only 1.7% of ELCs created during the study period were designated for rice production (Table S18), suggesting that rice production and price were not driving forces behind ELC allocation. The negative relationship between the producer price of cassava and new ELC allocation was strong (two-year time lag rate ratio = 0.982, Figure 4). Cassava is not a valuable crop, yet it was the third most designated crop for new ELCs during the study period (4.9% of new ELCs, Table S18). It is unclear what is driving the negative relationship between cassava and new ELCs after two years.

**Table 3. Parameter coefficients, standard errors, and rate ratios from the top model(s) in the macroeconomic analysis with rate of economic land concession allocation response. Missing values denote predictor variables that were not selected in the top model(s) for that lag period.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***No time lag*** | | |  | ***1 year time lag*** | | |  | ***2 year time lag*** | | |
| **Variable** | **Coefficient** | **SE** | **Rate ratioa** |  | **Coefficient** | **SE** | **Rate ratioa** |  | **Coefficient** | **SE** | **Rate ratioa** |
| ***Macroeconomic*** |  |  |  |  |  |  |  |  |  |  |  |
| GDP | - | - | - |  | -0.01500 | 0.00340 | 0.985 |  | -0.02600\* | 0.00390 | 0.974 |
| Agricultural proportion of GDP | 0.27000 | 0.07000 | 1.310 |  | 0.25000 | 0.06600 | 1.284 |  | -0.03400\* | 0.07600 | 0.967 |
| Development flows - agriculture | - | - | - |  | - | - | - |  | -0.00005\* | 0.00020 | 1.000 |
| Development flows - environment | 0.03100 | 0.00400 | 1.031 |  | - | - | - |  | -0.00260\* | 0.00450 | 0.997 |
| Foreign direct investment | - | - | - |  | 0.00360 | 0.00050 | 1.004 |  | 0.00040\* | 0.00060 | 1.000 |
| Population density | -4.43000 | 0.85000 | 0.012 |  | -6.09000 | 0.81000 | 0.002 |  | -7.68000\* | 0.95000 | 0.000 |
| Forest remaining | -0.00030 | 0.00004 | 1.000 |  | -0.00004 | 0.00004 | 1.000 |  | 0.00004\* | 0.00005 | 1.000 |
| ***Commodity / production*** |  |  |  |  |  |  |  |  |  |  |  |
| Change in median market price - corn | -0.00330 | 0.005697 | 0.997 |  | 0.00704\* | 0.00647 | 1.007 |  | -0.00365\* | 0.00329 | 0.996 |
| Change in median market price - rice | 0.009324 | 0.00198 | 1.009 |  | -0.00429\* | 0.00272 | 0.996 |  | 0.00004\* | 0.00058 | 1.000 |
| Change in median market price - rubber | 0.001247 | 0.00024 | 1.001 |  | 0.00019\* | 0.00022 | 1.000 |  | -0.00004\* | 0.00009 | 1.000 |
| Change in median market price - sugar | -0.00005 | 0.001931 | 1.00 |  | 0.00708\* | 0.00127 | 1.007 |  | 0.00877\* | 0.00124 | 1.009 |
| Non-food agricultural production index | -0.00995 | 0.00175 | 0.990 |  | 0.00672\* | 0.00264 | 1.007 |  | -0.00149\* | 0.00203 | 0.999 |
| Crop production index | - | - | - |  | 0.00042\* | 0.00144 | 1.000 |  | -0.00328\* | 0.00427 | 0.997 |
| Total production from forestry | - | - | - |  | 0.00000\* | 0.00000 | 1.000 |  | 0.00000\* | 0.00000 | 1.000 |
| Forest remaining | -0.00014 | 0.00002 | 1.000 |  | -0.00017\* | 0.00003 | 1.000 |  | -0.00013\* | 0.00003 | 1.000 |
| ***Producer prices*** |  |  |  |  |  |  |  |  |  |  |  |
| Producer price of corn | 0.00415 | 0.00355 | 1.004 |  | 0.01093\* | 0.00240 | 1.011 |  | 0.00014\* | 0.00081 | 1.000 |
| Producer price of rice | -0.02465 | 0.00436 | 0.976 |  | 0.00452\* | 0.00564 | 1.005 |  | 0.01258\* | 0.00474 | 1.013 |
| Producer price of rubber | 0.03424 | 0.00401 | 1.035 |  | -0.00075\* | 0.00228 | 0.999 |  | -0.00431\* | 0.00467 | 0.996 |
| Producer price of sugar | 0.00004 | 0.00010 | 1.000 |  | 0.00016\* | 0.00018 | 1.000 |  | 0.00000\* | 0.00006 | 1.000 |
| Producer price of cassava | 0.00032 | 0.00123 | 1.000 |  | 0.00006\* | 0.00076 | 1.000 |  | -0.01791\* | 0.00214 | 0.982 |
| Forest remaining | -0.00023 | 0.00002 | 1.000 |  | -0.00015\* | 0.00002 | 1.000 |  | -0.00013\* | 0.00002 | 1.000 |

\* Coefficients derived from full averaging of models within dAIC < 6.

A Rate ratio = exp(coefficient)

Chart

Description automatically generated

**Figure 2. Modelled relationships between economic predictors and the allocation of new economic land concessions in Cambodia between 1993 – 2015. Top row: no time lag between predictor and response; middle row: 1-year time lag between predictor and response; bottom row: 2-year time lag between predictor and response.**

A picture containing diagram

Description automatically generated

**Figure 3. Modelled relationships between commodity price predictors and the allocation of new economic land concessions in Cambodia between 1993 – 2015. Top two rows: no time lag between predictor and response; third row: 1-year time lag between predictor and response; bottom row: 2-year time lag between predictor and response.**

Diagram, engineering drawing

Description automatically generated

**Figure 4. Modelled relationships between producer price predictors and the allocation of new economic land concessions Cambodia between 1993 – 2015. Top row: no time lag between predictor and response; middle row: 1-year time lag between predictor and response; bottom row: 2-year time lag between predictor and response.**

***Socioeconomic analysis***

*Commune-level model*

Initial within-set model selection resulted in a final candidate set with 10 models and 13 unique variables (Table S19). There was a single top model according to AIC (m1), with all other models having delta AIC values of more than 18 (Table S19). The top model only had one non-control variable - population density (Table 4). The random effects term with the highest variance was Commune (10.45 [SD = 3.23], 60% of the total random effect variance), followed by Province (6.77 [SD = 2.60], 39% of the total random effect variance, Table 4). The variance explained by year at both the commune and province level was low (0.005 [SD = 0.068] and 0.0005 [SD = 0.022] respectively), contributing approximately 1% of the total random effect variance (Table 4). The marginal R2 (fixed effects only) was 0.78 (78%), and the conditional R2 (fixed and random effects) was 1, suggesting that most of the model variance was explained by the fixed effects. The largest positive effect was from mean elevation (rate ratio = 2.861, Table 4) which relates to 0.6 forest pixels (0.06 km2) predicted within an “average” commune (i.e., all other fixed and random effects set to their mean) when mean elevation is at the minimum within the country. When the mean elevation is at the maximum found within the country (and all other terms are set to their mean), the number of forest pixels predicted is 13,380 (1,204 km2). This highlights that higher elevation areas of Cambodia are much more likely to be forested than lower elevation areas. The strongest negative effect was from population density (rate ratio = 0.001, Table 4) which relates to approximately 1.5 predicted forest pixels (0.14 km2) at the minimum value of population density found within the country, contrasting with a prediction of effectively zero (2.22 × 10-16) forest pixels at the highest value of population density within the country. All other model terms, excluding the presence of ELCs, had positive effects on forest cover (Table 4). These effects suggest that remote communes (large distances to provincial capitals) that are centrally located within the country (far away from international borders) are predicted to have high forest cover. Interestingly, although the effects are weak, communes that contain ELCs are predicted to have lower forest cover than those without, and communes with protected areas are predicted to have higher forest cover than those without (Table 4).

The results from the final commune-level model must, however, be viewed with extreme caution because model validation revealed some serious underlying issues. As is suggested by the variance associated with the commune-level random effect term, there was extreme variation between communes for all variables (predictors and response, Figure 5). This between-group variance results in the model being unsuitable for generalised (i.e., ‘global’) predictions (Figure 5). Intercept and slope estimates between communes, even within the same province, varied hugely (Figure 6), and this issue was highlighted in diagnostic plots where we see that the assumption of normality of deviations of the conditional means of the random effects (for commune) from the global intercept is violated (Figure S6). Furthermore, the model residuals displayed heteroskedasticity, with the model predicting particularly poorly for lower values of the response (Figure S7). Therefore, drawing general inferences about the relationships between forest cover and socioeconomics at the country level using this model is inappropriate.

**Table 4. Model outputs and rate ratios from the top models from the socioeconomic analysis. Outputs are for the commune-level analysis and the province-level analysis. Reported coefficients are on the link (log) scale.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Variance** | **Std.Dev** | **Coefficient** | | **SE** | **Rate ratio*a*** |
| ***Commune-level final model*** |  |  | |  |  |  |
| *Random effects* |  |  |  | |  |  |
| Commune (intercept) | 10.4500 | 3.2334 | - | | - |  |
| Year/Commune (slope) | 0.0046 | 0.0680 | - | | - |  |
| Province (intercept) | 6.7730 | 2.6025 | - | | - |  |
| Year/Province (slope) | 0.0005 | 0.0220 | - | | - |  |
| *Fixed effects* |  |  |  | |  |  |
| Intercept | - | - | -4.6240 | | 0.5620 |  |
| Population density | - | - | -7.5140 | | 1.1270 | 0.001 |
| Mean elevation | - | - | 1.0510 | | 0.1220 | 2.861 |
| Distance to In'tl border | - | - | 0.5805 | | 0.2036 | 1.787 |
| Distance to Provincial capital | - | - | 0.6929 | | 0.1114 | 2.000 |
| ELC presence | - | - | 0.0000 | | 0.0025 | 0.999 |
| PA presence | - | - | 0.0093 | | 0.0143 | 1.009 |
| ***Province-level final model*** |  |  |  | |  |  |
| *Random effects* |  |  |  | |  |  |
| Province (intercept) | 1.1762 | 1.0845 | - | | - |  |
| Year/Province (slope) | 0.0058 | 0.0765 | - | | - |  |
| *Fixed effects* |  |  | - | | - |  |
| Intercept | - | - | -2.9900 | | 0.4497 |  |
| Males in school (low) | - | - | 0.0051 | | 0.0019 | 1.002 |
| Distance to school (low) | - | - | -0.0174 | | 0.0022 | 1.002 |
| Mean elevation (low) | - | - | -0.0223 | | 0.0024 | 1.002 |
| Distance to border (low) | - | - | 0.0061 | | 0.0019 | 1.002 |
| Distance to Prov capital (low) | - | - | -0.0072 | | 0.0019 | 1.002 |
| Presence of economic concessions (1) | - | - | 1.9974 | | 0.4090 | 1.505 |
| Presence of PAs (1) | - | - | 2.8063 | | 0.4965 | 1.643 |

*a* Rate ratio = exp(coefficient)

Chart, scatter chart

Description automatically generated

**Figure 5. Predicted relationships (red lines) between socioeconomic variables and forest cover in Cambodia between 2007 – 2012 from the top commune-level model. Predictions are ‘global’ i.e., all random effects were set to their mean values, and thus predictions are not for any specific commune. Black dots are the raw data points of each predictor versus forest cover.**

Diagram, shape, arrow

Description automatically generated

**Figure 6. Predicted relationships between population density and forest cover within Cambodian provinces between 2007 – 2012 using the top commune-level model. Faded grey lines are the predictions for each individual commune within each province. Black lines are the mean provincial predictions, which were computed using the 50% quantile from all commune predictions. Plot panels have non-standard y axis ranges.**

*Province-level model*

The province-level models were run to eliminate the commune-level variation and to identify any broader relationships between forest cover and socioeconomics. A candidate set of 19 models was built and an evaluation of AIC selected a single model (m8) as the top model (Table S20). Model m5 had some support (delta AIC = 5, Table S20) but was a simpler version of m8 and therefore inferences were drawn from m8 alone. The random effects term with the highest variance was Province (1.18 [SD = 1.08], which constituted 99% of the total random effects variance), followed by year (0.006 [SD = 0.077], which was 1% of the total random effects variance). The marginal R2 (fixed effects only) was 0.71 (71%) and the conditional R2 (fixed and random effects) was 0.99 (99%), suggesting that the majority of model variance was explained by the fixed effects. Presence of ELCs and presence of PAs had the largest two positive effects relative to their refences levels (no ELCs, no PAs), suggesting that provinces that have those two features are predicted to also have higher forest cover (rate ratios = 1.51 and 1.64 respectively). In provinces where the proportion of males in school and distance to school are both low, higher levels of forest cover are predicted compared with provinces where these variables are high. Furthermore, in provinces where elevation, distance to an international border, and distance to the provincial capital are low, forest cover is predicted to be higher than in provinces where these variables are high. However, all the above effects are weak (Figures 7 & 8). For example, the difference in the predicted number of forest pixels between a province with a low proportion of males in school and a province with a high proportion (with all other variables set to low), is 200 (18 km2). The difference in the number of predicted forest pixels between a province with low median distances to schools and a province with high median distances (with all other variables set to low), is 689 (62 km2). As standalone figures these appear large, but in the context of the range of the response variable (minimum value of 54 forest pixels to a maximum of 146,876 forest pixels), the effects are relatively weak. Presence of PAs had the largest effect on predicted forest pixels. The number of forest pixels predicted for a province with PA presence is 36,890 (3,320 km2) higher than for a province with no PA presence. This emphasises the relationship between forested land and protected areas in Cambodia. The size of the effects for the two socioeconomic predictors (proportion of males in school, and distance to school) in the top model suggest that these variables have little power to predict forest cover at the provincial level in Cambodia, but that the presence of protected areas and economic land concessions do.

***Cluster analysis***

The UPGMA clustering had the highest cophenetic correlation (0.79) and the lowest Gower distance (254.14) and was therefore selected. The matrix correlation statistic suggested that 4 clusters were optimal, but that between 3 and 7 clusters had very similar support. When divided by 4 clusters, there was a large group (16) of provinces that fell into a single cluster, and so 5 clusters were chosen to add further nuance (Figure 9). The provinces within clusters were geographically contiguous (Figure 10), although clusters that had smaller cophenetic distances (i.e., were closer on the dendrogram, Figure 9) were not necessarily geographically contiguous. The largest cluster (cluster 5) dominated a central strip of the country, separating the smaller, and more similar clusters (Figure 10). Only clusters 2 and 4 were contiguous with each other. These results suggest that provinces often have similar socioeconomic conditions to that of their neighbours, but that there are also distinct regions within the country that can be characterised by their socioeconomics rather than their geography. A heatmap of the socioeconomic variable values for each cluster revealed some distinguishing patterns (Figure 11). The largest cluster (cluster 5) was distinguished by high or very high values of all variables, which translates to generally large provinces with high population density, high education levels, high proportions of primary and secondary sector workers, and high migration (Table 5). This contrasts with cluster 2, which has predominantly low values for the socioeconomic variables which translates to very small provinces with low population density, low levels of education, low levels of primary sector employment (higher secondary sector employment), and low levels of migration (Table 5). Clusters 3 and 4 had the highest levels of migration (and interestingly the highest levels of land conflict), education, and population density, reflecting the presence of two of the three largest cities and significant urban development. Cluster 1 had the lowest population density, education, proportion of secondary sector workers, and migration, reflecting the clusters remote geography and rural character. Provinces within cluster 1 were also the most forested but had also lost the most forest during the study period (Figure 12). Provinces within cluster 5 were generally the next most forested after cluster 1 and had also lost large areas of forest during the study period (Figure 12). Cluster 3 had the least amount of forest, which was expected due to high levels of urbanisation and agriculture. Clusters 1 and 2 had the highest elevation, and clusters 1 and 5 had the highest mean distance to a provincial capital (Figure 12).

A picture containing diagram

Description automatically generated

**Figure 7. Predicted forest cover within each Cambodian province given high and low levels of school attendance (males aged 6 – 24 in school) from the top province-level model. All other variables in the model were set to their reference level (distance to school = low, elevation = low, distance to international border = low, distance to provincial capital = low, economic land concession = yes, protected area = yes).**

A picture containing diagram

Description automatically generated

**Figure 8. Predicted forest cover within each Cambodian province given high and low distances to the nearest school from the top province-level model. All other variables in the model were set to their reference level (school attendance = low, elevation = low, distance to international border = low, distance to provincial capital = low, economic land concession = yes, protected area = yes).**

Chart

Description automatically generated

**Figure 9. Cambodian provinces clustered based on socioeconomics. Data were averaged across the study period 2007 – 2012. Variables included were total population, population density, number of land conflict cases, number of criminal cases per capita, number of in- and out-migrants, the proportion of the population classified as indigenous, proportion of males aged 6 – 24 in school, proportion of the population employed in the primary and secondary sectors, proportion of families with no access to agricultural land, proportion of families who kept pigs, distance to the nearest school, proportion of families with access to waste collection, and distance to the commune (administrative) centre. The clustering method was unweighted pair-group using arithmetic averages (UPGMA).**

Map

Description automatically generated

**Figure 10. Map of Cambodia showing the clusters resulting from the unweighted pair-group using arithmetic averages (UPGMA) method. Provinces are labelled. The upper white polygon is the Tonle Sap lake, and the lower white polygon is the city of Phnom Penh, both of which were excluded from the analysis.**

Chart, bar chart

Description automatically generated

**Figure 11. Heatmap showing the variable values for each cluster. Variables were categorised as “v.low” if the mean (across provinces within that cluster) was below the 25% quantile for that variable across the whole country, “low” if the mean was above 25 and below 50%, “high” if the mean was above 50% but below 75%, and “v.high” if the mean was above the 75% quantile. Pax\_migt\_out = numbers of out-migrants, Pax\_migt\_in = numbers of in-migrants, land\_confl = number of land conflicts, crim\_case = criminal cases per capita, KM\_Comm = distance to commune office, garbage = proportion of families with access to waste collection, dist\_school = distance to nearest school, pig\_fam = proportion of families who keep pigs, Les1\_R\_Land = proportion of families with no rice land, propSecSec = proportion of adults employed in the secondary sector, propPrimSec = proportion of adults employed in the primary sector, M6\_24\_sch = proportion of males aged 6-24 in education, prop\_ind = proportion of the population that is indigenous, pop\_den = population density.**

Diagram

Description automatically generated

**Figure 12. Boxplots showing the distribution of environmental variables for each cluster: *a* = mean forest area, *b* = mean area (km2), *c* = change in forest cover (between 2007-2012), *d* = mean elevation (masl), *e* = mean distance to international border, *f* = mean distance to a provincial capital. Boxplots show the median (centre line within boxes), 25 and 75% percentiles (box edges), and minimum and maximum values (upper and lower whiskers, not exceeding 1.5 × interquartile range). 5 UPGMA clusters.**

**Table 5. Descriptive typology of the provinces and clusters within Cambodia, clustered using socioeconomic variables and the unweighted pair group using arithmetic mean (UPGMA)**

|  |  |  |
| --- | --- | --- |
| **UPGMA cluster** | **Provinces** | **Description** |
| 1 | Mondulkiri, Ratanikiri | Very large provinces with very high elevations. Very low population density, and very high proportion of indigenous people. Very low education levels, very high proportion of primary sector workers and very low proportion of secondary sector workers. Economic security provided by rural livelihoods - few people have no farmland and livestock ownership is common. Very low access to services, high crime per capita, low land conflict, and very low migration levels. |
| 2 | Pailin | Very small province with very high elevations. Low population density and low proportion of indigenous people. Low levels of education, low proportion of people in the primary sector but higher proportion of people in the secondary sector. Very few people with no farmland, but very little livestock ownership. High access to services and high crime per capita. Low land conflict and low migration. |
| 3 | Kampong Cham, Kandal, Prey Veng, Takeo | Small provinces with very low elevations. Very high population density and high proportion of indigenous people. Very high levels of education, high proportion of people in the primary sector, but very high proportion of people in the secondary sector. High proportion of people with no farmland, but high levels of livestock ownership. High access to services and low crime per capita. But very high migration levels and very high rates of land conflict. |
| 4 | Banteay Meanchey, Battambang | Large provinces with low elevations. Very high population density and very low proportion of indigenous people. Very high levels of education, and relatively low proportion of workers in the primary and secondary sectors (suggesting higher proportions in the other sectors e.g., tertiary). High proportion of people with no farmland, and low levels of livestock ownership (suggesting very urban). Low access to services, but this may be explained by the mean size of the provinces in this cluster (there is high access to garbage collection). Low crime per capita, but very high migration and very high rates of land conflict |
| 5 | Kampong Chhnang, Kampong Speu, Kampong Thom, Kampot, Kep, Koh Kong, Kracheh, Otdar Meanchey, Preah Sihanouk, Preah Vihear, Pursat, Siem Reap, Stung Treng, Svay Rieng | Large provinces with high elevations. High population density and very high proportion of indigenous people. High levels of education, and a high proportion of workers in both primary and secondary sectors. Very high proportion of people with no farmland, but also very high proportion of people with livestock. Low access to services (although very high access to garbage collection) - this may be an artefact of the very large mean area of the provinces in this cluster. Very high crime rates, very high migration, and very high rates of land conflict. |

**Discussion**

Understanding the drivers and proximate causes of forest cover loss and land use change is critical for the development of sustainable environmental policies and forest conservation initiatives. Studies need to target multiple scales to build a cohesive picture of the social-ecological systems within which deforestation occurs, so that policy development is appropriate and effective. Importantly, researchers need to select the appropriate method to answer specific questions at specific scales, and the complexity of the system must be understood. In this study, we have modelled the relationships between macroeconomic and socioeconomic variables and forest loss, forest cover, and the development of new ELCs. We have investigated these relationships at multiple scales using a variety of approaches and have revealed some important relationships from which we can make cautious inferences regarding direct and indirect drivers of forest loss. Furthermore, we have revealed key methodological issues, particularly around scale and model variance, that are likely to be common in these types of analyses, but which often remain unexplored or unreported in the literature.

*Macroeconomic analysis*

There were very few significant effects in the macroeconomic analysis with forest loss as the response variable. Some of the predictor variables have been shown to correlate with LUC in other studies, such as GDP (Ewers, 2006; Fan and Ding, 2016; Gong et al., 2013; Kuang et al., 2016), the contribution of economic sectors to national GDP (Gong et al., 2013), human population growth and density (Bonilla-Bedoya et al., 2018; Fan and Ding, 2016), and agricultural output (Fan and Ding, 2016). There are several possible explanations for the lack of effects in this study. First, previous studies have been at different scales to this study, such as global (e.g., Ewers, 2006), or sub-national (e.g., Gong et al., 2013), and therefore the drivers which are operating at those scales may be different to the drivers operating at the national scale in Cambodia. Second, Cambodia’s economy is unique within Asia because of the civil unrest and war, economic collapse, and subsequent rapid economic revival. This may render comparison of macroeconomic drivers of forest loss and LUC with other Asian countries ineffective. For example, Cambodia’s economy is in its infancy relative to many other countries in the region, and therefore forest loss during the study period may have been driven more by local drivers such as poverty, insecure land tenure, and land speculation by migrants, rather than national-level economics. Third, we did not include predictor variables covering institutional factors, land rights or tenure, or environmental policies, which have been shown to be important (Culas, 2007). Fourth, we only investigated up to two years of time lag between changes in predictor variables and changes in forest cover. It is possible that the effects of macroeconomics on forest cover and LUC operate at a larger temporal scale than considered in this study.

In contrast, the macroeconomic variables were effective at predicting the allocation of new ELCs. Although ELCs do not guarantee deforestation (indeed not all ELCs are awarded on forested land), the deforestation rates within ELCs are up to 105% higher than comparable areas with no ELCs (Davis et al., 2015). There has also been widespread allocation of ELCs within forested community land and protected areas, resulting in the loss of important forest habitat, rural livelihoods, and indigenous land rights (Global Witness, 2013; Watson et al., 2014). Therefore, ELCs themselves can be considered direct drivers of forest loss, rendering the macroeconomic predictors indirect drivers. Our results have demonstrated that during the study period, the economic development of the country was closely linked to the increase in ELCs, which in turn have driven forest loss. There were clear relationships between the size of the agricultural sector, the rates of foreign investment, and the number of new ELCs. The process of awarding ELC contracts in Cambodia has been criticised for lacking transparency and for corruption (Neef et al., 2013; Vrieze and Kuch, 2012), and so it is not always possible to identify who owns a particular concession. Nevertheless, of those identified, 48% were foreign owned (Licadho, 2019). Despite real and perceived benefits of attracting foreign investment and expanding the production of cash crops, there are numerous negative effects on local people and the environment. Development of potential agricultural land by investors comes with opportunity costs for local people, who otherwise may have had access to the land, water, and other resources, and could have developed agriculture themselves that would have alleviated poverty more effectively than a(n externally owned) commercial agricultural enterprise. Furthermore, agro-industrial production of cash crops for international markets leaves the country open to price shocks and other suboptimal market fluctuations (De Schutter, 2011).

Changes in new ELC allocation can also be effectively predicted by several key agricultural commodity prices, both on the international market and internally at the farmgate scale. Rubber, sugar, corn, and rice we all important variables in the models, and increases in the market prices of these commodities can predict increases in the allocation of ELCs. Importantly, there were differences in the effects of commodity and producer prices on ELC allocation at different time lags, suggesting that either investors will delay investing in a new crop for up to two years after the prices increase, or that the process of purchasing land and establishing an ELC venture can in some cases be a slow process. International market forces are known to drive LUC, and globally, land conversion for commodity production is the single largest driver of deforestation (Curtis et al., 2018). Grogan et al (2015) provide an empirical example of how the international market price of rubber can drive deforestation in frontier areas of Cambodia and Vietnam. Understanding which commodities are driving land conversion, the strength of the effects, the time lags, and the legal and institutional mechanisms that facilitate the link between prices and forest loss, is critical for predicting future forest loss. The Cambodian Prime Minister issued a moratorium on new ELCs in 2012, which drastically reduced (although did not eliminate) ELC allocation. Although this has had a positive effect on forests, rural livelihoods, and indigenous land tenure, it is unclear how long this reprieve will last (Davis et al., 2015), or whether a new mechanism will emerge to replace ELCs. The opaque legal mechanisms and weak institutional frameworks that governed ELCs in the past are likely to continue to hinder the development of sustainable agricultural policies. The relationships between macroeconomics, commodity prices, and industrial-scale agriculture identified in this study are likely replicated across the region, and therefore these results will be of use to researchers and policy makers outside Cambodia.

*Socioeconomic analysis*

The commune-level model revealed that population density was the only non-control variable with any effect, and the effect was very weak. We were limited in the socioeconomic variables that were available, and it is possible that the variables selected were simply poor predictors of forest cover. However, the modelling process revealed very large between-commune variation in both predictor and response variables, in addition to a large number of random effect levels (between 1,317 and 1,512). Model predictions from the final model, and from preliminary models, showed that the parameter estimates (intercepts and slopes) for a given socioeconomic variable (see Figure 6 for an example from population density) varied widely from commune to commune, even within the same province. Therefore, it is possible that the difficulty in estimating a single parameter from the surrounding “noise” resulted in the detection of weak, or no effects, rather than a genuine lack of effects. The province-level model was built to counter the issue of excessive between-commune variance by approaching the analysis from a different scale. Two socioeconomic variables remained in the final province-level model but again, the effects were relatively weak. It is still possible that the weak effects represented a genuine lack of correlation between socioeconomics and forest cover, however, modelling the effects at a larger scale will simply mask the large variation that exists at the finer scale, rather than eliminating it. This analysis highlights the importance of scale when modelling complex social-ecological systems; researchers must not only select the scale of the analysis carefully but must also be aware of underlying variation which may be affecting estimates, requiring cautious interpretation of results. The results of the socioeconomic analysis have further highlighted the effect of scale on drivers with larger effects. The direction of the effect of the presence of ELCs changes depending on whether you are looking at the commune-level or the province-level (ELCs have a negative effect on forest cover within a commune, but a positive effect on forest cover within a province). This reversal of effect direction between scales also occurs for distance to an international border (positive within communes and negative within provinces) and the distance to the provincial capital (positive within communes and negative within provinces). Taken together, the two models can add important nuance to the interpretation of results; provinces that are small and are close to international borders have higher forest cover, but within those provinces, the communes that are furthest away from the border and the provincial capital are predicted to have the highest forest cover. These results demonstrate how the relationships between forest cover and predictor variables are being driven in different directions at different scales, emphasising the complexity of modelling social-ecological systems.

The inherent complexity within social-ecological systems results in significant challenges when researchers attempt to model them (Basse et al., 2014). Taking this study as an example, a researcher has a choice between modelling at a large scale (e.g., national, regional) where effects may be weak or unrepresentative of much of the country or region, or modelling at a fine scale where effects may be swamped by variation resulting in the loss of the true signal. In some cases, prudence may stop researchers gathering increasingly complex data, but rather reframe their analytical goal by removing hypothesis testing and aiming for description of the data rather than explanation. Advances in simulation modelling and machine learning can isolate our thinking and increase understanding, without the need for large datasets and complex statistical modelling procedures (refs).

*Cluster analysis*

In this study we investigated the use of cluster analysis to describe Cambodia in terms of socioeconomics. The analysis revealed interesting patterns of distinct regions, suggesting that in many cases provinces that are adjacent to each other tend to have similar socioeconomic characteristics, resulting in clusters that are spatially contiguous. The two cluster that generally display the largest differences are clusters 1 and 3. Cluster 1 contains the provinces of Mondul Kiri and Rattank Kiri which are large, remote, and some of the least developed provinces in the country. They are home to the Eastern Plains Landscape which is one of the most important areas in SEA for biodiversity (Chapter 2, Gray et al., 2012; Griffin and Nuttall, 2020; Nuttall et al., 2017). The cluster has the highest forest cover, low population density, low access to services, and low migration. Economic development in the first two decades after the civil war was focused almost entirely on the major cities: Phnom Penh (cluster 3), Sihanoukville (cluster 5), and Battambang (cluster 4), with rural provinces remaining underdeveloped, inaccessible, and poor (Hughes and Un, 2011). The lack of infrastructure and access, coupled with low population density and few employment opportunities that limited in-migration, has meant that forest cover has remained high (Evans et al., 2013). Conversely, cluster 3 has the lowest levels of forest cover and contains the capital city of Phnom Penh and the surrounding provinces which are the hubs for industry and economic activity (such as the garment sector). Cluster 5 is interesting because it contains the largest number of provinces. The expectation was that the provinces that most closely resembled cluster 1 (i.e., large, rural provinces with high forest cover) such as Stung Treng, Preah Vihear, and Koh Kong, would have been clustered either with cluster 1, or within a separate cluster. However, they were clustered with the central belt of provinces (e.g., Kampong Speu, Kampong Chhnang, Kampong Thom) which are almost exclusively low elevation agricultural provinces that are geared towards rice production. The inclusion of Stung Treng, Preah Vihear, and Koh Kong within this cluster and the resulting cluster typologies, suggest that there has been some success in increasing the socioeconomic status of rural, highly forested provinces without excessive loss of forest cover.

The advantage of clustering techniques such as UPGMA is that although there are metrics that can suggest optimal numbers of clusters, the researcher can select the number of clusters that is most useful for their particular investigation (Borcard et al., 2018). Unlike statistical models, cluster analysis does not produce estimates of effect sizes, nor can predictions be made. Nevertheless, by altering the number of clusters, investigating different clustering approaches, followed by considered exploratory analysis and plotting, a comprehensive picture of the study system can be produced. This may be a sensible first step in a larger analysis which can increase understanding of the system before modelling approaches are decided upon. Furthermore, methods such as cluster analysis are conceptually simpler than advanced statistical and mechanistic modelling, making interpretation and explanation to non-specialist audiences, such as policy makers, simpler.

*Conclusion*

Cambodia is a country rich in natural resources and biological diversity, and despite the targeted efforts of the Khmer Rouge regime, is also rich in cultural and social diversity. Once the economic powerhouse of Indochina, in the 1960s Cambodia was the world’s third largest exporter of milled rice, behind only Thailand and the United States (Hughes and Un, 2011). It is therefore appropriate that political leaders are given recognition for bringing the economy from complete collapse during the civil war, to a growth rate in 2006 that was larger than any other Asian economy apart from China (Solcomb, 2010). Over the last two decades there has been significant improvements in access to services, poverty, and inequality, thanks to pro-poor growth in consumption, which together pushed Cambodia’s poverty reduction well beyond the Millennium Development Goal targets (World Bank, 2014). However, relative metrics of inequality (e.g., Gini Index) mask the actual gap between the rich and the poor in absolute terms, which has been increasing dramatically (World Bank, 2014). There exists a very large wealth gap between urban and rural populations, and between the urban rich and urban poor, and the gaps are growing (Solcomb, 2010). Some of the economic mechanisms which have vastly increased the wealth of the urban political class, whilst violating local land rights and driving deforestation, such as ELCs, have been justifiably criticized (Davis et al., 2015; Global Witness, 2013; Vrieze and Kuch, 2012). Further land use policies such as Directive 01 and social land concessions have lacked transparency, have been poorly implemented, and have eroded protected forests (Thesis appendix, Grimsditch and Schoenberger, 2015; Milne, 2013).

*Conclusions. What forest transition pathway is Cambodia on (lambin & Meyfroidt papers)? Linking forest cover and forest loss to economic and social factors is challenging. This is particularly true of a country like Cambodia which has changed and developed extremely rapidly over the last 30 years – does not conform to development trends of the region. The governance of the country also means that natural resource exploitation is difficult to pin down through official metrics – opaque legal mechanisms and processes (e.g. ELC allocation). Nevertheless, there are analytical tools to use. Which ones you choose will depend on the data you have and the questions you are trying to ask. Final point of optimism - something about the Environmental Kuznet curve and hoping that Cambodia can reduce forest loss before its too late.*

**References**

Abdullah, S.A., Nakagoshi, N., 2007. Forest fragmentation and its correlation to human land use change in the state of Selangor, peninsular Malaysia. For. Ecol. Manag. 241, 39–48. https://doi.org/10.1016/j.foreco.2006.12.016

Aguiar, A.P.D., Vieira, I.C.G., Assis, T.O., Dalla-Nora, E.L., Toledo, P.M., Oliveira Santos-Junior, R.A., Batistella, M., Coelho, A.S., Savaget, E.K., Aragão, L.E.O.C., Nobre, C.A., Ometto, J.P.H., 2016. Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. Glob. Change Biol. 22, 1821–1840. https://doi.org/10.1111/gcb.13134

Aspinall, R., 2004. Modelling land use change with generalized linear models—a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. J. Environ. Manage., Modelling land use change and environmental impact 72, 91–103. https://doi.org/10.1016/j.jenvman.2004.02.009

Barrett, C.B., Gibson, C.C., Hoffman, B., McCUBBINS, M.D., 2006. The Complex Links between Governance and Biodiversity. Conserv. Biol. 20, 1358–1366. https://doi.org/10.1111/j.1523-1739.2006.00521.x

Bartoń, K., 2020. MuMIn: Multi-Model Inference.

Basse, R.M., Omrani, H., Charif, O., Gerber, P., Bódis, K., 2014. Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. Appl. Geogr. 53, 160–171. https://doi.org/10.1016/j.apgeog.2014.06.016

Beever, E.A., Simberloff, D., Crowley, S.L., Al‐Chokhachy, R., Jackson, H.A., Petersen, S.L., 2019. Social–ecological mismatches create conservation challenges in introduced species management. Front. Ecol. Environ. 17, 117–125. https://doi.org/10.1002/fee.2000

Bhattarai, M., Hammig, M., 2004. Governance, economic policy, and the environmental Kuznets curve for natural tropical forests. Environ. Dev. Econ. 9, 367–382. https://doi.org/10.1017/S1355770X03001293

Bonilla-Bedoya, S., Estrella-Bastidas, A., Molina, J.R., Herrera, M.Á., 2018. Socioecological system and potential deforestation in Western Amazon forest landscapes. Sci. Total Environ. 644, 1044–1055. https://doi.org/10.1016/j.scitotenv.2018.07.028

Borcard, D., Gillet, F., Legendre, P., 2018. Cluster Analysis, in: Borcard, D., Gillet, F., Legendre, P. (Eds.), Numerical Ecology with R, Use R! Springer International Publishing, Cham, pp. 59–150. https://doi.org/10.1007/978-3-319-71404-2\_4

Brown, D.G., Verburg, P.H., Pontius, R.G., Lange, M.D., 2013. Opportunities to improve impact, integration, and evaluation of land change models. Curr. Opin. Environ. Sustain., Human settlements and industrial systems 5, 452–457. https://doi.org/10.1016/j.cosust.2013.07.012

Burnham, K.P., Anderson, D.R., 2007. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer Science & Business Media.

Ceddia, M.G., 2019. The impact of income, land, and wealth inequality on agricultural expansion in Latin America. Proc. Natl. Acad. Sci. 116, 2527–2532. https://doi.org/10.1073/pnas.1814894116

Ceddia, M.G., Gunter, U., Corriveau-Bourque, A., 2015. Land tenure and agricultural expansion in Latin America: The role of Indigenous Peoples’ and local communities’ forest rights. Glob. Environ. Change 35, 316–322. https://doi.org/10.1016/j.gloenvcha.2015.09.010

Chapman, P.M., Tobias, J.A., Edwards, D.P., Davies, R.G., 2018. Contrasting impacts of land-use change on phylogenetic and functional diversity of tropical forest birds. J. Appl. Ecol. 55, 1604–1614. https://doi.org/10.1111/1365-2664.13073

Chhair, S., Ung, L., 2013. Economic history of industrialization in Cambodia (Working Paper No. No.7), WIDER. World Institue for Development Economics Research.

Culas, R.J., 2012. REDD and forest transition: Tunneling through the environmental Kuznets curve. Ecol. Econ. 79, 44–51. https://doi.org/10.1016/j.ecolecon.2012.04.015

Culas, R.J., 2007. Deforestation and the environmental Kuznets curve: An institutional perspective. Ecol. Econ. 61, 429–437. https://doi.org/10.1016/j.ecolecon.2006.03.014

Curtis, P.G., Slay, C.M., Harris, N.L., Tyukavina, A., Hansen, M.C., 2018. Classifying drivers of global forest loss. Science 361, 1108–1111. https://doi.org/10.1126/science.aau3445

Davis, K.F., Yu, K., Rulli, M.C., Pichdara, L., D’Odorico, P., 2015. Accelerated deforestation driven by large-scale land acquisitions in Cambodia. Nat. Geosci. 8, 772–775. https://doi.org/10.1038/ngeo2540

Dawson, T.P., Rounsevell, M.D.A., Kluv\’ankov\’a-Oravsk\’a Tatiana, Chobotov\’a Veronika, Stirling, A., 2010. Dynamic properties of complex adaptive ecosystems: implications for the sustainability of service provision. Biodivers. Conserv. 19, 2843–2853. https://doi.org/10.1007/s10531-010-9892-z"

De Schutter, O., 2011. How not to think of land-grabbing: three critiques of large-scale investments in farmland. J. Peasant Stud. 38, 249–279. https://doi.org/10.1080/03066150.2011.559008

de Souza, R.A., De Marco, P., 2014. The use of species distribution models to predict the spatial distribution of deforestation in the western Brazilian Amazon. Ecol. Model. 291, 250–259. https://doi.org/10.1016/j.ecolmodel.2014.07.007

Eliste, P., Zorya, S., 2015. Cambodian agriculture in transition: Opportunities and risks. World Bank Group, Washington DC, USA.

Estoque, R.C., Ooba, M., Avitabile, V., Hijioka, Y., DasGupta, R., Togawa, T., Murayama, Y., 2019. The future of Southeast Asia’s forests. Nat. Commun. 10, 1829. https://doi.org/10.1038/s41467-019-09646-4

Evans, T., O’Kelly, H., Men, S., Nut, M., Pet, P., Pheakdey, P., Pollard, E., 2013. Seima Protection Forest, in: Evidence-Based Conservation: Lessons from the Lower Mekong. Routledge, London, UK, pp. 157–185.

Ewers, R.M., 2006. Interaction effects between economic development and forest cover determine deforestation rates. Glob. Environ. Change 16, 161–169. https://doi.org/10.1016/j.gloenvcha.2005.12.001

Fan, Q., Ding, S., 2016. Landscape pattern changes at a county scale: A case study in Fengqiu, Henan Province, China from 1990 to 2013. CATENA 137, 152–160. https://doi.org/10.1016/j.catena.2015.09.012

Frewer, T., Chan, S., 2014. GIS and the ‘Usual Suspects’-[Mis]understanding Land Use Change in Cambodia. Hum. Ecol. 42, 267–281. https://doi.org/10.1007/s10745-013-9639-z

Gao, J., Liu, Y., 2012. Deforestation in Heilongjiang Province of China, 1896–2000: Severity, spatiotemporal patterns and causes. Appl. Geogr. 35, 345–352. https://doi.org/10.1016/j.apgeog.2012.08.001

Gaughan, A.E., Binford, M.W., Southworth, J., 2009. Tourism, forest conversion, and land transformations in the Angkor basin, Cambodia. Appl. Geogr. 29, 212–223. https://doi.org/10.1016/j.apgeog.2008.09.007

Geist, H.J., Lambin, E.F., 2002. Proximate Causes and Underlying Driving Forces of Tropical DeforestationTropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. BioScience 52, 143–150. https://doi.org/10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2

Gelman, A., Hill, J., 2006. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Global Witness, 2013. Rubber Barons: How Vietnamese companies and international financiers are driving a land grabbing crisis in Cambodia and Laos. Global Witness.

Gong, C., Yu, S., Joesting, H., Chen, J., 2013. Determining socioeconomic drivers of urban forest fragmentation with historical remote sensing images. Landsc. Urban Plan. 117, 57–65. https://doi.org/10.1016/j.landurbplan.2013.04.009

Gray, T.N.E., Prum, S., Pin, C., Phan, C., 2012. Distance sampling reveals Cambodia’s Eastern Plains Landscape supports the largest global population of the Endangered banteng Bos javanicus. Oryx 46, 563–566. https://doi.org/10.1017/S0030605312000567

Griffin, O., Nuttall, M., 2020. Status of Key Speceis in Keo Seima Wildlife Sanctuary 2010-2020 (Status report). Wildlife Conservation Society, Phnom Penh, Cambodia.

Grimsditch, M., Schoenberger, L., 2015. New actions and existing policies: The implementation and impacts of Order 01. NGO Forum.

Grogan, K., Pflugmacher, D., Hostert, P., Kennedy, R., Fensholt, R., 2015. Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series. Remote Sens. Environ. 169, 438–453. https://doi.org/10.1016/j.rse.2015.03.001

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J., Inger, R., 2018. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 6, e4794. https://doi.org/10.7717/peerj.4794

Hearn, A.J., Cushman, S.A., Goossens, B., Macdonald, E., Ross, J., Hunter, L.T.B., Abram, N.K., Macdonald, D.W., 2018. Evaluating scenarios of landscape change for Sunda clouded leopard connectivity in a human dominated landscape. Biol. Conserv. 222, 232–240. https://doi.org/10.1016/j.biocon.2018.04.016

Hughes, A.C., 2018. Have Indo-Malaysian forests reached the end of the road? Biol. Conserv. 223, 129–137. https://doi.org/10.1016/j.biocon.2018.04.029

Hughes, A.C., 2017. Understanding the drivers of Southeast Asian biodiversity loss. Ecosphere 8, e01624. https://doi.org/10.1002/ecs2.1624

Hughes, C., Un, K., 2011. Cambodia’s Economic Transformation (No. 49), NIAS Studies in Asian Topic. Nordic Institute of Asian Studies, Leifsgade 33, DK-2300 Copenhagen S, Denmark.

Imai, N., Furukawa, T., Tsujino, R., Kitamura, S., Yumoto, T., 2018. Factors affecting forest area change in Southeast Asia during 1980-2010. PLOS ONE 13, e0197391. https://doi.org/10.1371/journal.pone.0197391

Kaimowitz, D., Fauné, A., 2003. Contras and Comandantes. J. Sustain. For. 16, 21–46. https://doi.org/10.1300/J091v16n03\_02

Kristensen, S.B.P., Busck, A.G., van der Sluis, T., Gaube, V., 2016. Patterns and drivers of farm-level land use change in selected European rural landscapes. Land Use Policy 57, 786–799. https://doi.org/10.1016/j.landusepol.2015.07.014

Kuang, W., Liu, J., Dong, J., Chi, W., Zhang, C., 2016. The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers. Landsc. Urban Plan. 145, 21–33. https://doi.org/10.1016/j.landurbplan.2015.10.001

Lambin, E.F., Meyfroidt, P., 2010. Land use transitions: Socio-ecological feedback versus socio-economic change. Land Use Policy, Forest transitions 27, 108–118. https://doi.org/10.1016/j.landusepol.2009.09.003

Licadho, 2019. Economic Land Concessions in Cambodia.

Luck, G.W., Smallbone, L.T., O’Brien, R., 2009. Socio-Economics and Vegetation Change in Urban Ecosystems: Patterns in Space and Time. Ecosystems 12, 604. https://doi.org/10.1007/s10021-009-9244-6

Mather, A.S., 1992. The Forest Transition. Area 24, 367–379.

Meyfroidt, P., Lambin, E.F., 2009. Forest transition in Vietnam and displacement of deforestation abroad. Proc. Natl. Acad. Sci. 106, 16139–16144. https://doi.org/10.1073/pnas.0904942106

Meyfroidt, P., Lambin, E.F., 2008. Forest transition in Vietnam and its environmental impacts. Glob. Change Biol. 14, 1319–1336. https://doi.org/10.1111/j.1365-2486.2008.01575.x

Milne, S., 2013. Under the leopard’s skin: Land commodification and the dilemmas of Indigenous communal title in upland Cambodia. Asia Pac. Viewp. 54, 323–339. https://doi.org/10.1111/apv.12027

Nakagawa, S., Schielzeth, H., 2017. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x@10.1111/(ISSN)2041-210X.STATSTOO

Neef, A., Touch, S., Chiengthong, J., 2013. The Politics and Ethics of Land Concessions in Rural Cambodia. J. Agric. Environ. Ethics 26, 1085–1103. https://doi.org/10.1007/s10806-013-9446-y

Nelson, G.C., Dobermann, A., Nakicenovic, N., O’Neill, B.C., 2006. Anthropogenic drivers of ecosystem change: An overview. Ecol. Soc. 11.

Nguyen, T.T., Do, T.L., Bühler, D., Hartje, R., Grote, U., 2015. Rural livelihoods and environmental resource dependence in Cambodia. Ecol. Econ. 120, 282–295. https://doi.org/10.1016/j.ecolecon.2015.11.001

Nuttall, M., Nut, M., Ung, V., O’kelly, H., 2017. Abundance estimates for the endangered Green Peafowl Pavo muticus in Cambodia: identification of a globally important site for conservation. Bird Conserv. Int. 27, 127–139. https://doi.org/10.1017/S0959270916000083

Poffenberger, M., 2006. People in the forest: community forestry experiences from Southeast Asia. Int. J. Environ. Sustain. Dev. 5, 57–69. https://doi.org/10.1504/IJESD.2006.008683

Price, S., 2020. War and Tropical Forests: Conservation in Areas of Armed Conflict. CRC Press.

QGIS Geographic Information System, 2020. . Open Source Geospatial Foundation Program.

R Core Team, 2021. R: A language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rudel, T.K., Coomes, O.T., Moran, E., Achard, F., Angelsen, A., Xu, J., Lambin, E., 2005. Forest transitions: towards a global understanding of land use change. Glob. Environ. Change 15, 23–31. https://doi.org/10.1016/j.gloenvcha.2004.11.001

Sodhi, N.S., Koh, L.P., Brook, B.W., Ng, P.K.L., 2004. Southeast Asian biodiversity: an impending disaster. Trends Ecol. Evol. 19, 654–660. https://doi.org/10.1016/j.tree.2004.09.006

Sodhi, N.S., Posa, M.R.C., Lee, T.M., Bickford, D., Koh, L.P., Brook, B.W., 2010. The state and conservation of Southeast Asian biodiversity. Biodivers. Conserv. 19, 317–328. https://doi.org/10.1007/s10531-009-9607-5

Solcomb, M., 2010. An economic histroy of Cambodia in the twentieth century. NUS Press, National University of Singapore.

Stern, D.I., 2004. The Rise and Fall of the Environmental Kuznets Curve. World Dev. 32, 1419–1439. https://doi.org/10.1016/j.worlddev.2004.03.004

Stevens, D., Dragićević, S., 2007. A GIS-Based Irregular Cellular Automata Model of Land-Use Change. Environ. Plan. B Plan. Des. 34, 708–724. https://doi.org/10.1068/b32098

Stibig, H.-J., Achard, F., Carboni, S., Raši, R., Miettinen, J., 2014. Change in tropical forest cover of Southeast Asia from 1990 to 2010. Biogeosciences 11, 247–258. https://doi.org/10.5194/bg-11-247-2014

Ty, T.V., Sunada, K., Ichikawa, Y., Oishi, S., 2012. Scenario-based Impact Assessment of Land Use/Cover and Climate Changes on Water Resources and Demand: A Case Study in the Srepok River Basin, Vietnam—Cambodia. Water Resour. Manag. 26, 1387–1407. https://doi.org/10.1007/s11269-011-9964-1

UNDP, 2020. Charting pathways out of multidimensional poverty: Achieving the SDGs. United Nations Development Program.

Vrieze, P., Kuch, N., 2012. Carving up Cambodia: One concesion at a time. Cambodia Dly. 4–11.

Watson, J.E.M., Dudley, N., Segan, D.B., Hockings, M., 2014. The performance and potential of protected areas. Nature 515, 67–73. https://doi.org/10.1038/nature13947

Wilcove, D.S., Giam, X., Edwards, D.P., Fisher, B., Koh, L.P., 2013. Navjot’s nightmare revisited: logging, agriculture, and biodiversity in Southeast Asia. Trends Ecol. Evol. 28, 531–540. https://doi.org/10.1016/j.tree.2013.04.005

World Bank, 2014. Where have all the poor gone? Cambodia poverty assessment 2013. World Bank Group, Washington DC, USA.

Yang, X., Zheng, X.-Q., Lv, L.-N., 2012. A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata. Ecol. Model. 233, 11–19. https://doi.org/10.1016/j.ecolmodel.2012.03.011

Zeng, Z., Estes, L., Ziegler, A.D., Chen, A., Searchinger, T., Hua, F., Guan, K., Jintrawet, A., Wood, E.F., 2018. Highland cropland expansion and forest loss in Southeast Asia in the twenty-first century. Nat. Geosci. 1. https://doi.org/10.1038/s41561-018-0166-9

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R, Statistics for Biology and Health. Springer Science+Business Media, New York, USA.