**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 themselves developed agriculture that would have alleviated poverty more effectively than a 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).

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, and 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 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 out of all the clusters. 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, suggests that there has been some success in increasing the socioeconomic status of rural, highly forested provinces without excessive loss of forest cover.

Cluster analysis. When the models above performed badly, we decided to try different approach. Specifically removing hypothesis testing – describe the data rather than trying to explain it. Not enough to just keep looking at increasingly complex data - simulations and machine learning can isolate our thinking and help to increase understanding. Talk about cluster results. What can be said about the country? How do the results relate to the modelling results? What are the similarities, what are the differences? Do we learn anything new?

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.