The Complex Crisis of Deforestation in Brazil: Ecology, Economy, and the Politics of Progress

CALEB WINEBRENNER

SSCI 574: Spatial Econometrics — Extra Credit Report
University of Southern California, Spatial Sciences Institute

September 21, 2021

Contents

Abstract

With the ecological impact of deforestation of the Amazon well-known, the question arises as to why such practices continue. This report looks at the problem through the lenses of urban economics and spatial econometrics to identify possible trends affecting the problem. In doing so, this analysis also explores the policies of current and past administrations on cutting forests, the role of indigenous peoples, and the available ecological data. Further, multiple regression models are explored to understand relationships between forest loss and other economic variables. Geographically-weighted regression holds the most potential for demonstrating clear relationships, though this approach is also not without its challenges. The results of these models are interpreted sociopolitically, including hypotheses for further study, new variables, and expected results. Finally, brief space is given to policy recommendations.

1. Introduction

The word *economy* comes originally from the Greek οἰκονομία, meaning "the management of the household" — in turn derived from the words οἶκος (*ecos*, house) and νόμος (*nomos*, law or custom).[*] By relation, it shares a root with the word *ecology*, this time combining οἶκος with the word for study, λογία (*logia*). This notion of economics evoking the well-being of a household or city is sensible, especially when considering the complex questions of land use being debated today. Likewise, the connection between economy and ecology is wise, for while there are considerations for humans in economic questions, there are also considerations of the well-being of the land and its nonhuman inhabitants.

As such, this research seeks to understand the implications for ecology and environmental justice (EJ) that are inextricably linked with questions human well-being. The relationship between two variables is explored: the Human Development Index for each state in Brazil, and the amount of forest cover in each state.

Theoretically speaking, this research is also framed by questions of urban economics, the notion of a "historically contingent" process of place, and the necessity of grassroots, participatory involvement for shaping policy.

2. Background

2.1 The Forests of Brazil

The forests of Brazil compose roughly 12% of the world's forest cover. Perhaps more important, those forests —- the Amazon being the best known —- provide the most important carbon sink on earth (helping to naturally control the amount of carbon dioxide and other greenhouse gasses in the atmosphere. In a word, Brazilian forests are critical for human well-being in the long term. Fig. 1 shows the percentage forest by state.

∗. Additionally, the word *econometrics* adds the word for "measure," μέτρον.

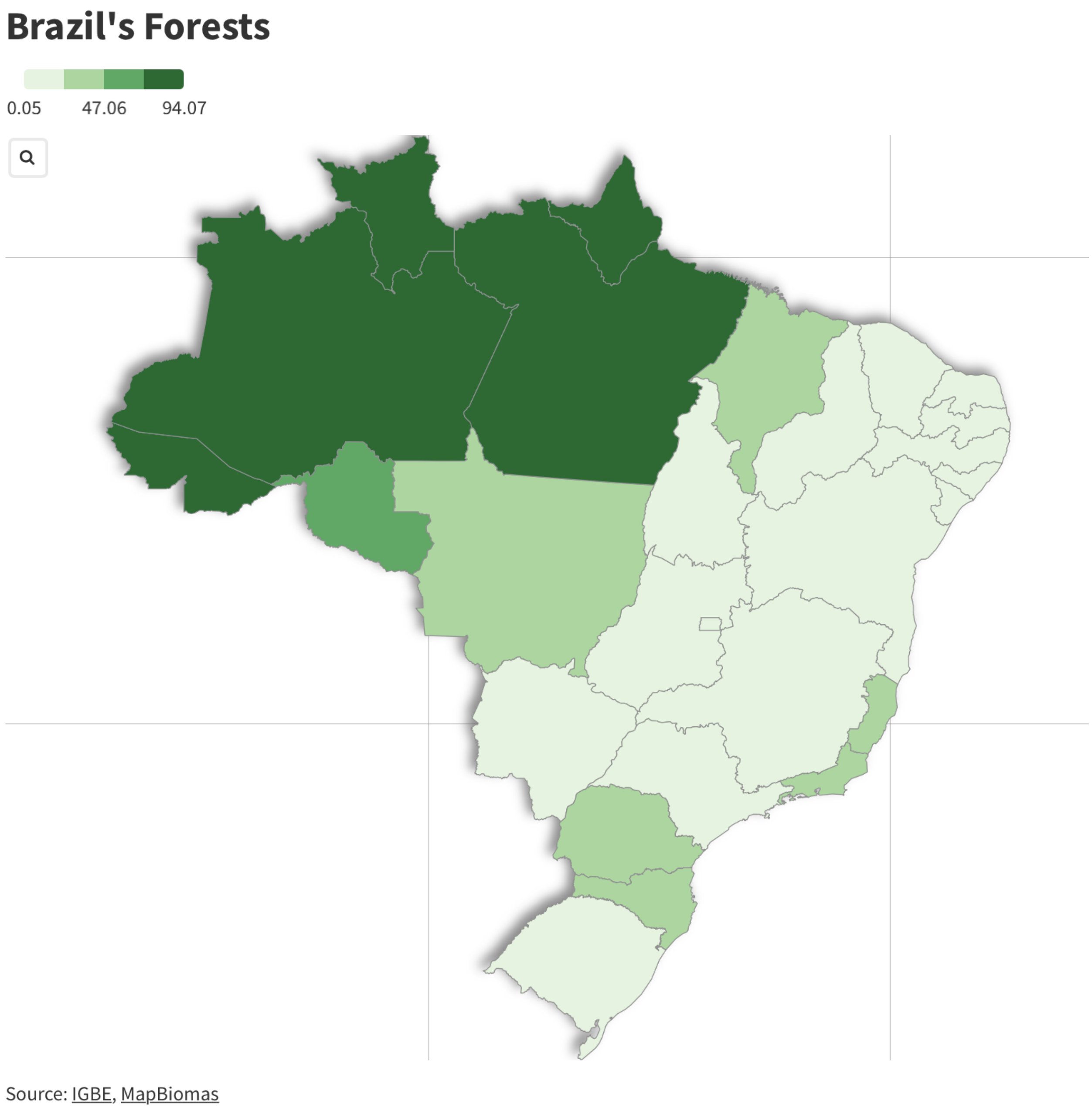

Figure 1: Percentage forest cover as of 2018.
Note: a dynamic version of this map is available on the project website.

2.2 The Political Climate

In 1992, The United Nations Sustainable Development Group met in Rio de Janeiro, Brazil. That meeting produced Agenda 21, a document outlining goals related to sustainable development of human settlements, health, and prosperity, with environmental protection and land management explicitly in mind. One such stated goal was to "establish systems for the assessment and systematic observations of forests and forest lands with a view to assessing the impact" of different kinds of human activity, such as settlement, agriculture, or commerce (Nations 1992, 11.30). The same document moreover mandated that any work toward sustainable development would be done while "Respecting the cultural integrity and the rights of indigenous people and their communities" and "Giving communities a large measure of participation in the sustainable management and protection of the local natural resources in order to enhance their productive capacity" (1992, 3.7). Despite Brazil being host to this conference, Brazil remains under close scrutiny for meeting these goals. While there are some protected lands and some recognized indigenous peoples, policies in Brazil remain contrary to the goals of protecting critical forests.

First, mid-twentieth century Brazil had policies of national development in the Amazon. The Amazon was viewed as a place of great potential wealth, where the forests could be used and the land transitioned to other uses. These policies set an explicit tone of slash-and-burn farming as an acceptible practice. Land use law even permits the clearing of land as a part of the process of claiming it (Cannon 2021). Conservation efforts in the 1960s in Brazil began in direct response to those policies.

In 2015, a law in Brazil was passed in their assembly that severely loosened any restrictions on the Amazon, effectively undoing gains by conservation groups. Thankfully, this law was vetoed and changes offered by the then-President of Brazil, Dilma Rousseff. The law still passed, but kept many restrictions and land-protections in place.

In 2019, President Jair Bolsano took office. Often viewed as a counterpart to Trump, Bolsano has been repeatedly criticized for turning a blind eye to illegal deforestation, even as he encourages agricultural business to legally clear the land for cattle ranching. While his policies have changed in early 2021 (see §7.), he nevertheless continued an ongoing trend.

To be clear, this is not to simply say that Brazilian policies have been without economic benefit. As we shall see (§*3.1*), the HDIin Brazil has been steadily increasing, and Brazil is continuing to play an important role in global economics.

2.3 Indigenous Rights

The Constitution of Brazil officially recognizes indigenous peoples as having a fundamental right to their cultures and practices. Indigenous tribes even have the right to file suit in courts. However, in practice this does not always translate to the kind of infrastructural support needed, including official representation — indigenous Brazilians are often left alone. Official data on these communities remains unclear, especially if comparing data from the Brazilian government and international organizations like the UN. Additionally, a recent report from the United Nations shows that the majority of the Amazon that is still intact is in fact held on indigenous lands (Cannon 2021). Indigenous peoples play an active role in organizing to protect the forest (such as the group Guardians of the Forest) and continue to exert pressure on official policy. As such, questions of indigenous lands remain a worthwhile consideration for this study (§6.).

2.4 Urban Economics

Scholarly research can also offer some insight into the growth of cities and the loss of forests. Pred (1984) offers a perspective from human geography on the arising of cities and the structuration of "space" into "place" by means of human institutions. More important, the author also explores the "transformation of nature" as one of the primary marks of human (and especially urban human) presence in a space, such as the felling of trees in order to build houses or develop farmlands. As the histories of a place emerge, Pred argues that these historical contingencies also solidify the agricultural and economic trends of a city, the patterns of land use, and the networks of cities that form in the landscape, including the patterns of social structure and employment. These patterns in the landscape, moreover, echo the kinds of development seen in cities under the von Thünen model, or the networks of settlements seen in central place theory. In short, Pred makes a place for social and political forces shaping human action. In this context, that leaves a place for environmental activism to be a counter-force to land development. Additionally, these overlapping theories offer a way to understand Brazil's highly urban south-central region as compared with the densely forested north.

Roos 2018 explicitly applies the von Thünen model to deforestation around urban centers, forest ecosystems, and the use of forests by a local population. More specifically, the study begins with an understanding of the competing interests of conservation groups, land developers, and the forest industry (i.e., companies that need the wood for materials). The paper integrates ecosystem services (ES) in an economic framework, wherein the services derived from an ecosystem are part of the local economy holistically, offering services and non-material benefits and not just a consumable resource. (I note too that cultural and social dimensions of ES mentioned in this paper, as well as the institutional forces that would shape ES, fit well with the Pred article above). However, the challenge lies in allocating the value of these services, especially against other potential uses of the land the the competing interests of different parties. Moreover, some of these benefits (such as green spaces increasing house value) can be understood economically, whereas other (such as higher reported well-being levels in communities with access to forests and green space) cannot be directly measured in economic terms. The von Thünen rings are introduced as a way to measure the distance and usefulness of a forest for local residents, in terms of social or cultural ES, as well as a way to model the decreasing value for residents the further into the hinterland that forests can be accessed (echoing the original bid rent curve). Additionally, the paper introduces service-dominant logic (SDL) as a way to valuate the services in question from cultural ES. In other words, the forest is seen as more than just the goods offered. In managing the various services available, policymakers, land developers, conservationists, and city residents can all play a role, benefiting from various services and affecting the institutions in place (see §7.). Finally, as urbanization continues, the framework here could be used to argue for both the economic and and ecological benefits of forests nearby to major urban centers.

3. Data

3.1 Data Table

Data sources are outlined in Table *3.1*.

Proposed Data			
Description	***Variable***	***Data Source***	***Use***
Forest Land	Hectares of land used	UN[a]	EDA
National HDI	Yearly HDI	UN[a]	EDA
Land use in Brazil	Deforestation	MapBiomas[b]	OLS/GWR
HDI by state	Indicators of Economic Wellbeing	Global Data Lab[c]	OLS/GWR

[a] Datasets from the UN are found at `http://data.un.org/Default.aspx`
These data are in `.csv` format and report the hectares of land devoted to a particular use.

[b] The Brazilian Annual Land Use and Land Cover Mapping Project, commonly known as MapBiomas, is a collective project of conservationists and GIS analysts in Brazil interested in gathering data on land use (especially forests) and mapping those changes over time. Data is made available for further analysis. The English-language statistics page can be found here:
`https://mapbiomas.org/en/estatisticas?cama_set_language=en`

[c] The Global Data Lab (`https://globaldatalab.org`) is a project of the Institute for Management Research at Radboud University in the Netherlands. Their datasets include measures of HDIat sub-national levels.

3.2 Fitness for Use

The data from the UN was a simple CSV, but only available at the national level. In this way, while it was good for exploratory data analysis (see §4.), it was limiting in what I could do for regression models. Hence, the data from Global Data Lab was used to provide HDIby state for each year.

The data from MapBiomas was more complicated. The project tracks numerous land changes, and as such this required that I make some critical choices in order to limit the scope of my project, such as eliminating mangrove forests from consideration (see §6.). Additionally, the data required extensive scrubbing in order to be used effectively. For the purposes of this study, I focused on the percent of forest cover per state.

3.3 Unexplored Variables

That said, it is also worth mentioning several variables that, given enough time, I would like to add to my study. These variables will help inform the discussion of my hypotheses and recommendations for policy, with the caveat that I have not actually run regressions with these variables. When I share the results (§5.), I also offer alternative variables to those listed above.

3.3.1 Amazônia Legal

The *Amazônia Legal*, or Brazil's Legal Amazon (BLA) is it called in English, is a socio-geographic region on Brazil. Fig. 2 shows the region, covering nine states of Brazil (though two only partially). It is worth noting that this region is designated for purposes of *economic* development. While conservation work seeks to influence that development policy, the region is not explicitly designated for ecological reasons. Nevertheless, given the nature of the study, the Amazon as an economic entity raises interesting questions. The BLA could be a function of the amount of a state's territory was in the region (i.e., the state of Amazonas, 1; Mato Grosso, 0.8; Rio Grande do Sul, 0). In parallel, the map below shows the two other economic regions, *Nordeste* (Northeast) and *Centro-Sul* (Central-South). Another option would be to include each region in regressions similar to the wieghting scheme used for the Legal Amazon (see §5. for more on this division).

Figure 2: Brazil's "Legal Amazon" — Source: Wikimedia Commons

3.3.2 Bioma Amazônia

While (BLA) is an economic entity, the Amazon *Biome* is an ecological one. This biome includes more than 200 areas that the Brazilian government protects on some level as a park, preserve, or similar. These lands, perhaps as a function of `Percentage-Land-Protected` by state, could be added to the regression.

4. Methods

In addition to the methods explored here, please see reports 2 & 3 for other methods used over the course of the semester. These reports will also include graphs from my exploratory data analysis. Additionally, these plots are available on the project website.

After exhausting the possibilities of other spatial regression methods, I returned to OLS to re-run my models, this time aiming to then use GWR. In short, OLS looks to see what sort of linear but aspatial relationship exists in the data. GWR then uses these models while creating a different set of coefficients for each areal unit observed. By mapping the coefficents and their *p*-values, GWR allows for a visual way to explore spatial relationships that are not clear in other spatial regressions.

Running the OLS regression in R produced the following output:

```
Call:
lm(formula = 'Percent-2008' ~ Health + Income + Education, data = df08)

Residuals:
    Min      1Q  Median      3Q     Max
-37.883 -11.626  -2.106  10.352  28.297

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -417.3      155.4  -2.684   0.0132 *
Health        1393.2      285.2   4.886 6.19e-05 ***
Income        -827.6      141.8  -5.838 5.99e-06 ***
Education     -149.4      203.0  -0.736   0.4692
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.53 on 23 degrees of freedom
Multiple R-squared:  0.6963,Adjusted R-squared:  0.6567
F-statistic: 17.58 on 3 and 23 DF,  p-value: 3.742e-06
```

Then OLS was also run for my 2018 data:

```
Call:
lm(formula = `Percent-2018` ~ Health + Income + Education, data = df18)

Residuals:
    Min      1Q  Median      3Q     Max
-32.630 -11.315  -0.845  10.733  30.193

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -229.19     187.73  -1.221  0.23451
Health       1042.08     303.48   3.434  0.00227 **
Income       -945.07     144.42  -6.544 1.12e-06 ***
Education      85.72     173.27   0.495  0.62548
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 17.3 on 23 degrees of freedom
Multiple R-squared:  0.6896,Adjusted R-squared:  0.6491
F-statistic: 17.03 on 3 and 23 DF,  p-value: 4.789e-06
```

These show that education is not statistically significant in the OLS.

Next, I began GWR(see the Appendix for my R code). Note that even though education was not statistically significant in the OLS, it was kept in the GWR. Several different methods were tried to find the best bandwidth and weighting scheme. Using a fixed kernel produced better results than the adaptive kernel, and the bi-square function proved the best weighting scheme:

```
Kernel function: gwr.bisquare
Fixed bandwidth: 18.50727
Summary of GWR coefficient estimates at data points:
                     Min.   1st Qu.    Median   3rd Qu.      Max.   Global
X.Intercept.    -743.429  -296.737   -94.913    11.576   434.722 -229.192
df18.Health     -287.227   384.917   553.359   832.979  1097.004 1042.082
df18.Income    -1319.358  -853.001  -539.226  -102.166   550.899 -945.074
df18.Education  -512.755  -234.091    33.541   203.027   775.899   85.722
Number of data points: 27
Effective number of parameters (residual: 2traceS - traceS'S): 12.7087
Effective degrees of freedom (residual: 2traceS - traceS'S): 14.2913
Sigma (residual: 2traceS - traceS'S): 10.3171
Effective number of parameters (model: traceS): 10.51059
```

```
Effective degrees of freedom (model: traceS): 16.48941
Sigma (model: traceS): 9.604856
Sigma (ML): 7.506051
AICc (GWR p. 61, eq 2.33; p. 96, eq. 4.21): 228.3694
AIC (GWR p. 96, eq. 4.22): 195.9816
Residual sum of squares: 1521.202
Quasi-global R2: 0.9313934
```

This model produced the lowest AIC and the highest Quasi-global R^2. Furthermore, an ANOVA test was run, demonstrating that the GWR fit the data better than OLS. Interestingly, when conducting the F3 test for spatial heterogeneity in the coefficients, education was statistically significant, but health was not. In the results section, I map the significance of income and education, based on the results of the F3 test.

5. Results

The main finding of the GWRwas that the relationship between the percentage of forest cover and income (as a variable within the HDI) was significant for the northern states in Brazil (see Fig. 3). Critically evaluating these results, however, could lead to a simple interpretation: that more urban, developed areas have higher income. In this sense, the variable of `Percent-Forest-Cover` does this research a disservice. As mentioned above, this is one place where different variables may lead to clearer results.

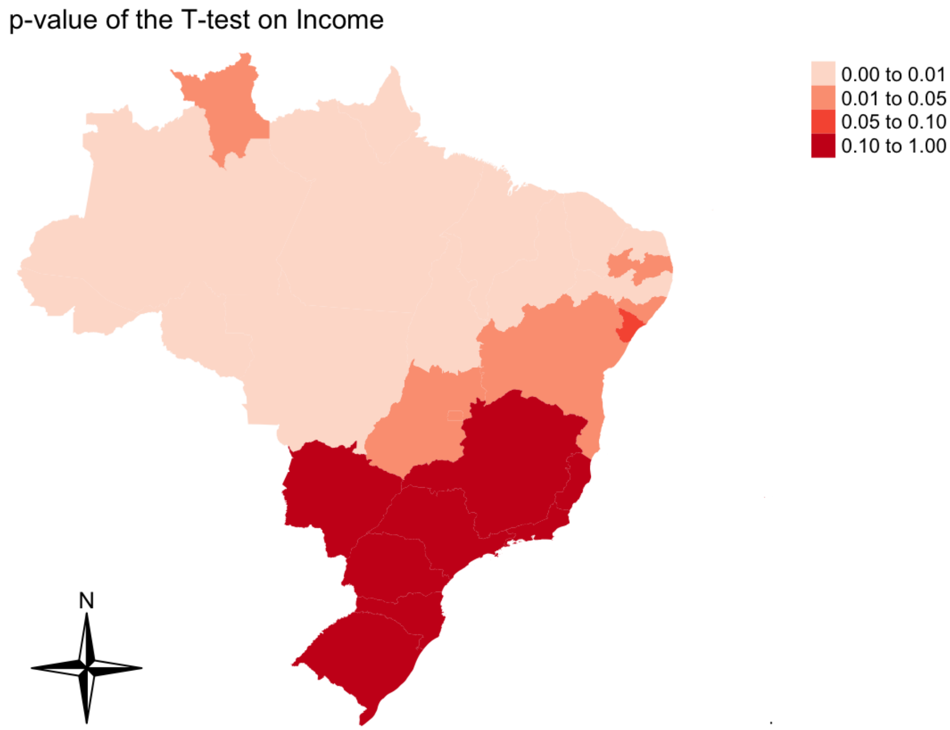

Figure 3: Mapping the significance of one of the GWR coefficients, Income

Additionally, if the GWR supports the conclusion of a North/South division in Brazil wherein the undeveloped north is seeing more deforestation as income increases, then further study would warrant looking at Brazil by its socio-geographic development regions, of which the Legal Amazon is only one. The Legal Amazon roughly corresponds to the lighter regions in the above map, while other shades roughly overlap with the other two development regions, thus indicating their usefulness for understanding these spatial relationships.

The relationship to education is more complicated (see Fig. 4). In the state of Amazonas (in the far northwest), which is also the state with the highest forest cover, the p-value is between 0.05 and 0.1 — so while possibly significant, it is only so at a higher threshold. On the other hand, highly urban states like Rio de Janeiro (in the southeast) don't show a strong relationship here at all. In other words, while education is indeed an important variable for the HDI, it doesn't seem to provide the same clarity as income does in understanding spatial relationships between the HDIand deforestation. As with the other variables, reconsidering these results in light of Brazil's development regions may prove insightful.

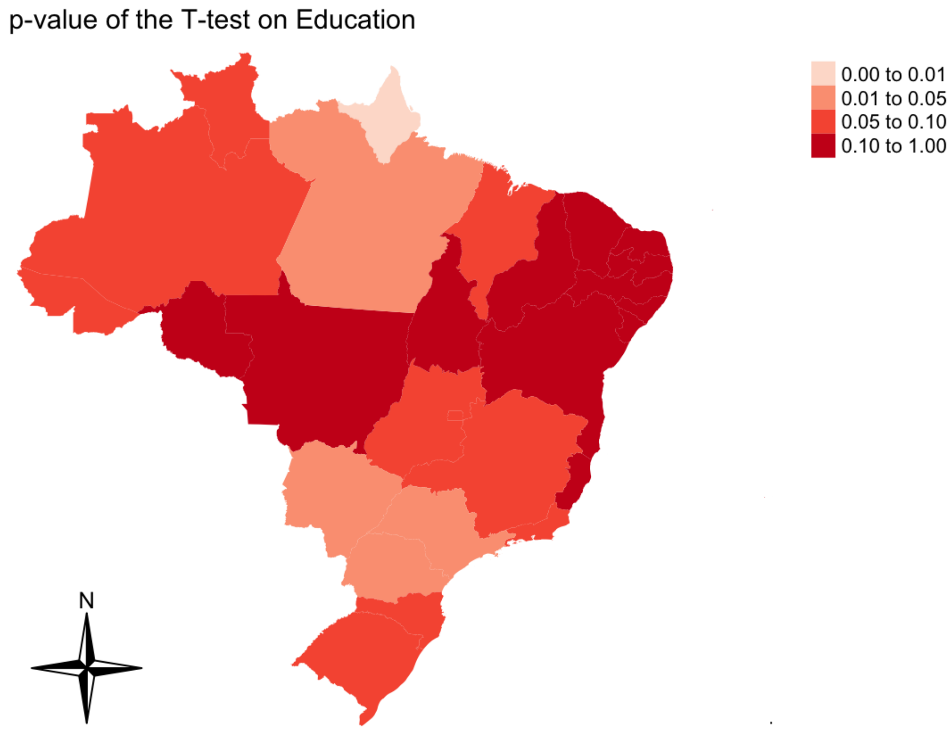

Figure 4: Mapping the significance of one of the GWR coefficients, Education

While it is clear to me that GWR offers a more nuanced way of approaching my questions, I remain unconvinced that my study used the best variables. In place of `Percent-Forest-Cover`, I'd also be interested in finding different variables for deforestation, such as accumulated forest loss. Such a variable could also include another critical type of forest, mangroves, as it's possible that limiting my study to one type of forest skewed my results.

6. Discussion

Several hypotheses emerged from this work. First, that the increase in HDI would lead to a loss of forests (i.e., not the other way around) because the policies currently in place encouraged forests to be cut in the name of economic growth. Secondly, that the results thus far, such as with income, were a result of urban growth in Brazil more broadly. In the service of these hypotheses, several lines of inquiry emerge for further study.

6.1 Remote Sensing

Within the last decade, groups such as Planet Labs have begun building "cubesats," or super-small satellites. This fleet of cubesats (roughly a one-meter cube), called "doves," results in high-resolution data *every 24 hours* (Planet Labs 2020). Previous problems with temporal resolution (whether expense or availability) are evaporating as technologies like these continue to be developed. What's more, in line with the Agenda 21, Planet Labs and companies like them create opportunities for public-private partnerships in the service of protecting forests. Therefore, given the issues with forest data, using remote sensing imagery analysis as a way to track changes may produce more reliable data.

It is worth noting, however, that such a step may also require some adjustments in study parameters, such as disaggregating the data to a municipal level or only monitoring a portion of Brazil, as image change analysis can be a labor- and computationally-intensive process.

As a side-note, remote sensing may also bolster my analyses in other ways. First, RS may be a way to empirically demonstrate the von Thünen model over time, as Brazilian cities develop and push back the hinterlands. Second, RS, along with fuzzy set theory, can be used the develop susceptibility maps (Bone et al. 2005). If applied to deforestation, such maps would provide a way to target interventions to prevent illegal deforestation.

6.2 Indigenous Lands

As stated previously, research is emerging that indigenous people are provably better stewards of their lands. Yet, without the needed economic and administrative support, indigenous peoples in Brazil could be limited in what they can do. A further line of inquiry, therefore, would be to include whether certain states or municipalities were on indigenous lands as a variable in spatial regressions. However, this comes with challenges around data availability. One possibility would be to use protected parks as a proxy for this, although not all protected lands are ones that overlap with lands held by indigenous peoples.

6.3 Local Spatial Knowledge

Furthermore, in addition to officially-reconized indigenous lands, local spatial knowledge (LSK) acquired by fieldwork and participatory mapping could be an approach for determining targeted areas on which to focus closely in monitoring the forest (Sletto et al. 2020). Such approaches could even influence policy by demonstrating how the forest is used sustainably. For example, consider the methods outlined by Delgado-Aguilar et al. 2019: participatory mapping techniques were used to determine places where an indigenous community in Ecuador gathered food, water, and supplies like firewood. These points were documented within a GIS. Remotely-sensed imagery of those forests were then input and a kernel density estimation performed, illustrating the relationships between forest degradation, human activity, and sustainable development (that is, not all uses of the forest caused degradation and the complex relationships lend themselves to further analysis and policy). A map from the study is shown in Fig. 5.

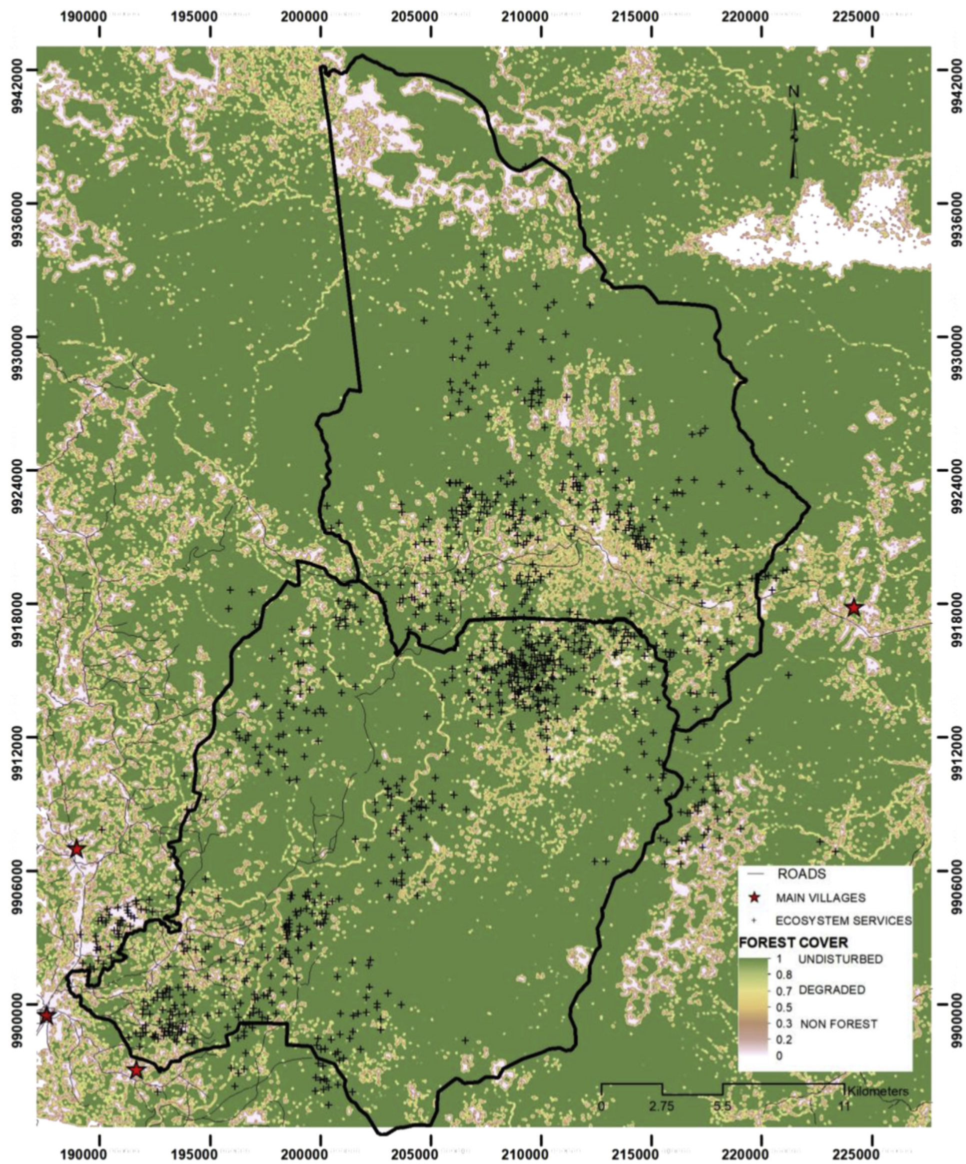

Figure 5: The Kichwe community in Ecuador and their use of the forests. Note the bold black line are areas of nationally protected land, ancestral to the tribe and limited in how they may be used. (Copied from Delgado-Aguilar et al. 2019)

7. Policy Recommendations

Just a few days ago, it was reported that Jair Bolsano has requested American aid in order to stop illegal deforestation. He also recently recommitted to being "deforestation-neutral" by 2030. While admirable on paper, a group of activists wrote to the Biden administration in April saying "Any project to help Brazil must be built from dialogue with civil society, subnational governments, academia, and, above all, with the local communities that know how to protect the forest and the goods and services it harbors" (quoted in Jones 2021). Given the nature of the Amazon, many of these local communities are the indigenous communities that are already working to protect their lands — the same ones already shown to be the best stewards of their lands.

In terms of policy, this could lead in several directions. Perhaps Bolsano's renewed call for climate action is a response to the new American administration — i.e., what's good for diplomacy is no longer just what is good for business. If so, this raises concerns about how much the Bolsano administration can be trusted, echoing the concerns of the activists quoted above. On the other hand, the last year has made the deforestation problem worse and so perhaps any aid is worth giving, even risking that such aid may be too little, too late.

The common denominator is appropriate oversight and data collection. Rather than work from estimated numbers, if academics could work with reliable information about the forest loss, more precise analyses and regressions would be possible. If policy could be built so as to have real-time monitoring of the forest using various remote sensing technologies, places for targeted interventions could be more easily determined. Finally, while indigenous communities have been shown to be reliable stewards of the forest, they nevertheless lack institutional funds and support.

Consider, for example, some of the work of Conservation International, a global nonprofit. One such upcoming project will be to train local and indigenous leaders in South America in the use of GIS technologies, in the service of mapping, preserving, and protecting their lands.[†] This project is intended to provide those communities with some of the skills needed for effective stewardship of their lands, even without official government support. Moreover, if the communities were to be provided with UAVs and trained to fly them, remotely-sensed data collection could also be a part of localized, grassroots efforts.

Returning to the theories outlined above, service-dominated logic (SDL) and ecosystem services may also have a role to play. If, under such a framework, the forests were not something to be used but a partner to be traded with, including providing ecotourism, then the economics of the situation could expand beyond profits for agricultural and logging companies. A more complex economic analysis, beyond my scope here, may prove beneficial.

Forest management is no longer a task of ecologists and GIS analysts alone; local communities, politicians, and policy-makers all play a role. The constitution of Ecuador, for example — which of note, is the site of the participatory study mentioned above — provides for the well-being of the environment as a constitutional right. To be clear, this isn't just for Ecuadorean citizens, but for the land and the forests themselves. Such a milestone in policy would not be possible without ongoing monitoring of the forests, whether by satellite, by drone, or by fieldwork from local communities. Were such a constitutional amendment passed in Brazil, perhaps the steady stream of data showing the ecological and economic impacts of deforestation can be put to use rebuilding the forests, not just preserving them.

†. I am unable to find a written source on this project. However, it is the focus of my summer internship, June to September 2021.

8. Conclusion

Given ample time and increased practice with `R`, I believe this study has a lot of promise. My remote sensing coursework this semester provided a great deal of useful literature on forest monitoring, but actually using remote sensing for data-gathering in this project was not yet feasible. Moreover, my upcoming summer internship, also devoted to ecological work with indigenous peoples in South America, will no doubt bolster my knowledge and skills in order to analyze the questions I've posted here.

Nevertheless, I have devoted what I can to the project over the course of the semester, and I've come to understand the problems of Brazilian deforestation in new ways. The data show that social and political factors play an enormous role in the life of the forests, even as such policies are made in the name of economic growth and human wellbeing. I've been able to quantify those relationships, even as qualitative hypotheses emerge. While it was frustrating at first to not have clear results and easily-mapped answers, I've come to be excited and curious at the prospect of continuing this project in the future, incorporating new variables and new lines of inquiry.

References

Bone, C., S. Dragicevic, and A. Roberts. 2005. "Integrating High Resolution Remote Sensing, GIS and Fuzzy Set Theory for Identifying Susceptibility Areas of Forest Insect Infestations." *International Journal of Remote Sensing* 26, no. 21 (November 10, 2005): 4809–4828.

Cannon, John C. 2021. "Nearly Half the Amazon's Intact Forest on Indigenous-Held Lands: Report." *Mongabay,* March 25, 2021.

Delgado-Aguilar, Ma. Jeaneth, Leonith Hinojosa, and Christine B. Schmitt. 2019. "Combining Remote Sensing Techniques and Participatory Mapping to Understand the Relations between Forest Degradation and Ecosystems Services in a Tropical Rainforest." *Applied Geography* 104 (March): 65–74.

Jones, Benji. 2021. "Deforestation in Brazil Is out of Control. Bolsonaro Is Asking for Billions to Stop It." *Vox,* May 1, 2021.

Nations, United. 1992. *Agenda 21.* Rio de Janeiro: United Nations Conference on Environment & Development.

Planet Labs. 2020. *Managing Forests Proactively with Satellite Imagery.* San Francisco: Planet Labs.

Pred, Allan. 1984. "Place as Historically Contingent Process: Structuration and the Time-Geography of Becoming Places." *Annals of the Association of American Geographers* 74 (2): 279–297. JSTOR: 2569284.

Roos, Anders, Jeannette Eggers, Cecilia Mark-Herbert, and Anders Lindhagen. 2018. "Using von Thünen Rings and Service-Dominant Logic in Balancing Forest Ecosystem Services." *Land Use Policy* 79 (December): 622–632.

Sletto, Bjørn, Alfredo Wagner Berno de Almeida, Joe Bryan, and Charles R. Hale, eds. 2020. *Radical Cartographies: Participatory Mapmaking from Latin America.* First edition. Austin: University of Texas Press.

Appendix: R code

````
**Note: Some of the blocks of `R` code here are copied from earlier
reports, in order to run additional analyses.**

# Introduction

```{r}
library(sp)
library(rgdal)
library(ggplot2)
library(GISTools)
library(spdep)
library(spatialreg)
library(RColorBrewer)
library(sf)
library(dplyr)
```
# Data

```{r}
df08 <- read.table(file = "data/2008-Data.csv", header = TRUE,
                    sep = ",", dec = ".", check.names = FALSE,
                    colClasses = c(`Percent-2008` = "numeric"))
df18 <- read.table(file = "data/2018-Data.csv", header = TRUE,
                   sep = ",", dec = ".", check.names = FALSE,
                   colClasses = c(`Percent-2018` = "numeric"))
```

# Methods
## EDA

Dr. Wu encouraged me to try a logarithmic transformation of the data.
This still resulted in a skewed histogram.
```{r}
hist(log(df18$`Percent-2018`))
m1 = lm(`HDI-2018` ~ log(`Percent-2018`), data = df18)
summary(m1)
```
````

ESDA

```{r}
## Reading the shapefile
brspatial <- readOGR(dsn = "data",
      layer="bra_admbnda_adm1_ibge_2020")
matrix <- poly2nb(brspatial, queen = TRUE)
head(brspatial)
```

```{r}
## Plotting the neighbor topography
plot(brspatial, border = "grey")
plot(matrix, coordinates(brspatial), add=TRUE, col="blue")
```

OLS & GWR

```{r}
library(sf)
library(tmap)
library(spgwr)
```

Given that any spatial relationships are proving hard to see, I decided to attempt other methods.
Specifically, I ran an OLS regression and then GWR.

```{r}
fit.ols.08<-lm(`Percent-2008`~
     `Health` + `Income` + `Education`, data = df08)
summary(fit.ols.08)
```
As before, education does not seem to be statistically significant, and will be discarded as a variable.

Now the same process, for the 2018 data:
```{r}
fit.ols.18<-lm(`Percent-2018`~
     `Health` + `Income` + `Education`, data = df18)
```

````
summary(fit.ols.18)
```
And now a GWR, with the bandwidth determined by the data.
```{r}
gwr.b1<-gwr.sel(df18$`Percent-2018` ~
     df18$Health + df18$Income, brspatial)
gwr.b1
```

```{r}
gwr.fit1<-gwr(df18$`Percent-2018` ~
     df18$Health + df18$Income + df18$Education,
data = brspatial, bandwidth = gwr.b1,
se.fit=T, hatmatrix=T)
gwr.fit1
```
In order to find a best model fit, I am also going to experiment with
different bandwidths and weighting schemes.

```{r}
gwr.b2<-gwr.sel(df18$`Percent-2018` ~
     df18$Health + df18$Income + df18$Education,
     data=brspatial, gweight = gwr.bisquare)
gwr.b2
```
With the new bandwidth, I'll try a second model:
```{r}
gwr.fit2<-gwr(df18$`Percent-2018` ~
     df18$Health + df18$Income + df18$Education,
data = brspatial, bandwidth = gwr.b2, gweight = gwr.bisquare,
se.fit=T, hatmatrix=T)
gwr.fit2
```

Since I want to find the best results, I will now try an
   adaptive kernel.
```{r}
gwr.b3<-gwr.sel(df18$`Percent-2018` ~
     df18$Health + df18$Income, data=brspatial, adapt = TRUE)
gwr.b3
````

````
gwr.fit3<-gwr(df18$`Percent-2018` ~
    df18$Health + df18$Income, data = brspatial,
    adapt=gwr.b3, se.fit=T, hatmatrix=T)
gwr.fit3
```

The adaptive kernel produced an error message.

First, examining the GWR models, the second model is the best,
   with an AIC of 195 and an R^2 of 93%. This model is what I will use!

Then several tests can be run in order to show that GWR is a better fit
   than OLS:
```{r}
BFC02.gwr.test(gwr.fit2)

BFC99.gwr.test(gwr.fit2)
```
The F3 test then allows us to examine which coefficients show
   spatial heterogeneity:
```{r}
LMZ.F3GWR.test(gwr.fit2)
```
This shows that for GWR, income and education are statistically
   significant in the coefficients.

```{r}
round(cor(as.data.frame(gwr.fit2$SDF[,2:11]), use ="complete.obs"),2)
```
Now, let's map the results

```{r}
names(gwr.fit2$SDF)
dfree<-gwr.fit2$results$edf
```
```{r}
## Calculating the t-stat for Income:
brspatial$Income.t <- gwr.fit2$SDF$df18.Income/gwr.fit2$SDF$df18.Income_se
brspatial$Income.t.p<-2*pt(-abs(brspatial$Income.t), dfree)
## Now, mapping the t-statistic:
````

```
breaks <- c(0,0.01,0.05,0.1,1)

tm_shape(brspatial, unit = "meter") +
  tm_polygons(col = "Income.t.p",palette = "Reds", breaks = breaks,
              border.alpha = 0, title = "") +
  tm_scale_bar(breaks = c(0, 1, 2), size = 1,
  position = c("right", "bottom")) +
  tm_compass(type = "4star", position = c("left", "bottom")) +
  tm_layout(main.title = "p-value of the T-test on Income",
  main.title.size = 0.95, frame = FALSE, legend.outside = TRUE)
```

```
```{r}
## Calculating the t-stat for Health:
brspatial$Health.t <- gwr.fit2$SDF$df18.Health/gwr.fit2$SDF$df18.Health_se
brspatial$Health.t.p<-2*pt(-abs(brspatial$Health.t), dfree)
## Now, mapping the t-statistic:
breaks <- c(0,0.01,0.05,0.1,1)

tm_shape(brspatial, unit = "meter") +
  tm_polygons(col = "Education.t.p",palette = "Reds", breaks = breaks,
              border.alpha = 0, title = "") +
  tm_scale_bar(breaks = c(0, 1, 2), size = 1,
  position = c("right", "bottom")) +
  tm_compass(type = "4star", position = c("left", "bottom")) +
  tm_layout(main.title = "p-value of the T-test on Health",
  main.title.size = 0.95, frame = FALSE, legend.outside = TRUE)
```
```