Spatial Economics – Assignment 4

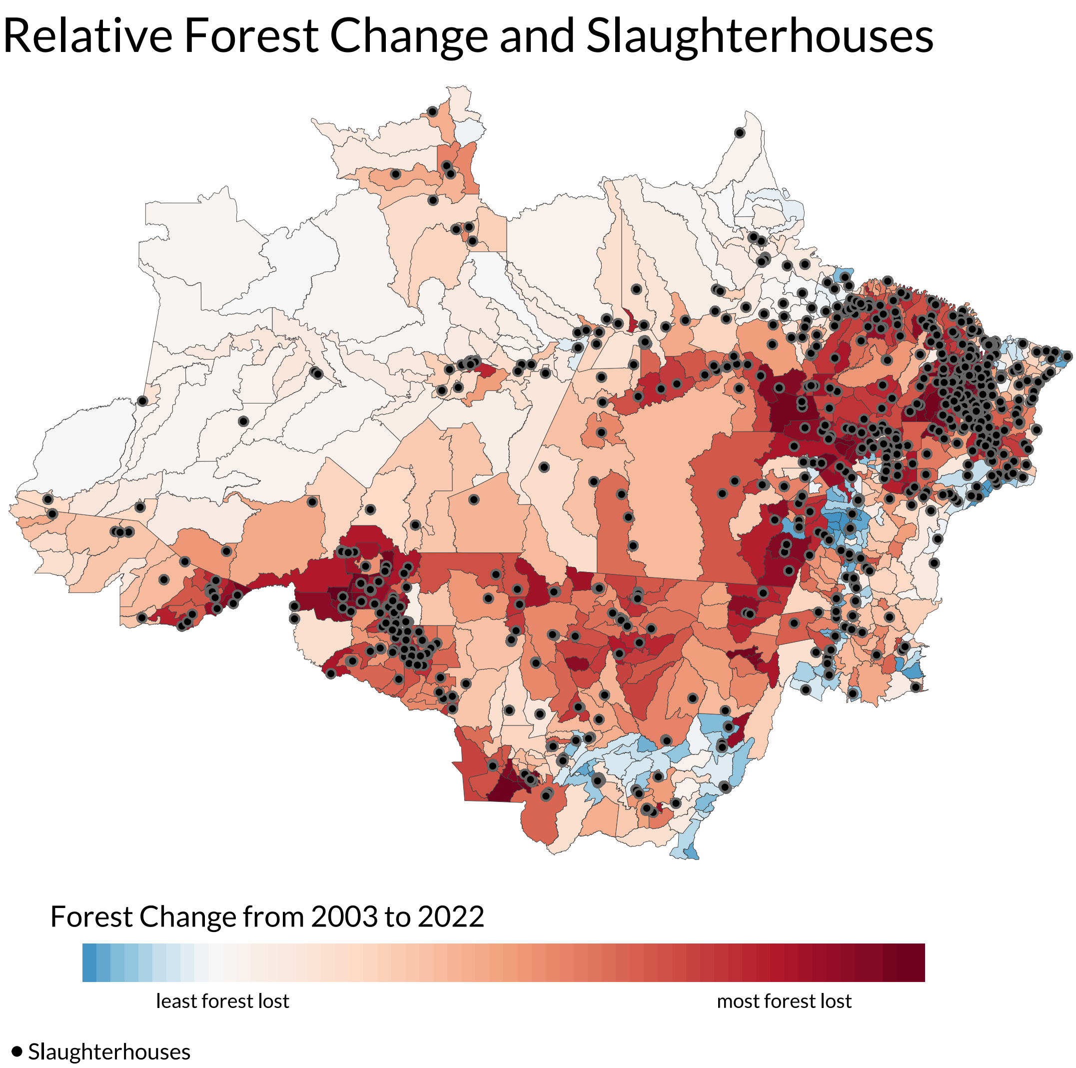
Max Heinze (h11742049@s.wu.ac.at)

June 21, 2024

Contents

The executable code that was used in compiling the assignment is available on GitHub at https://github.com/maxmheinze/spatial.

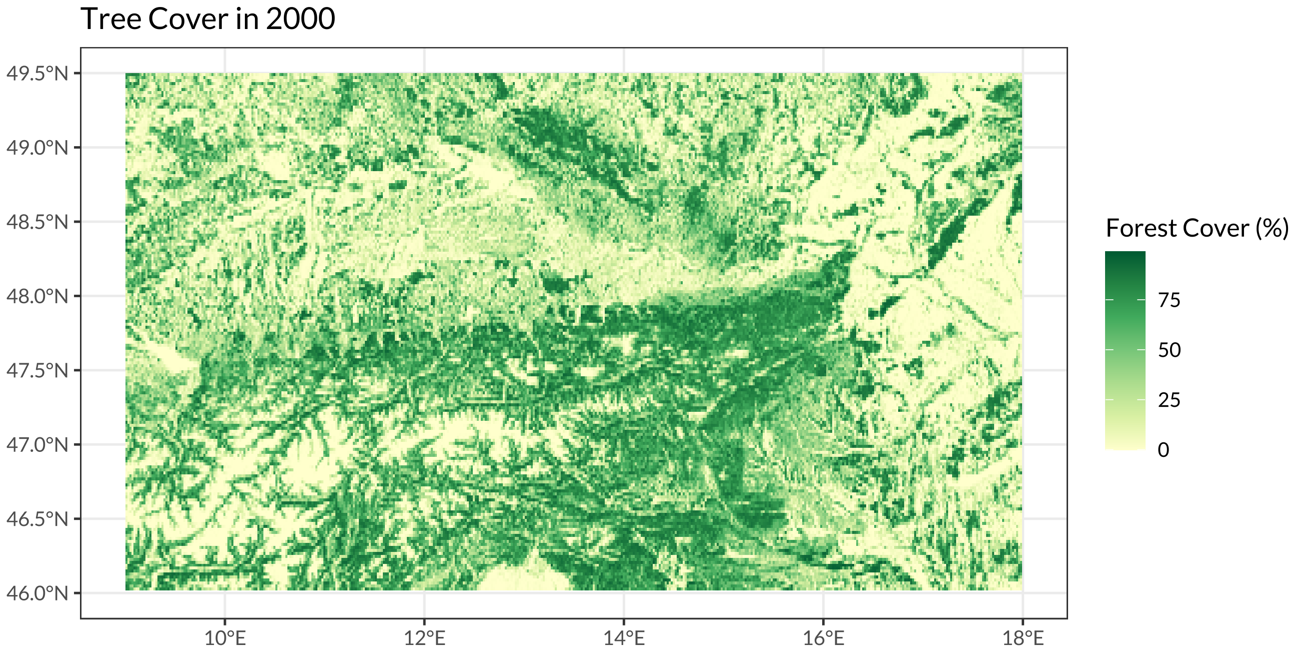
Task A



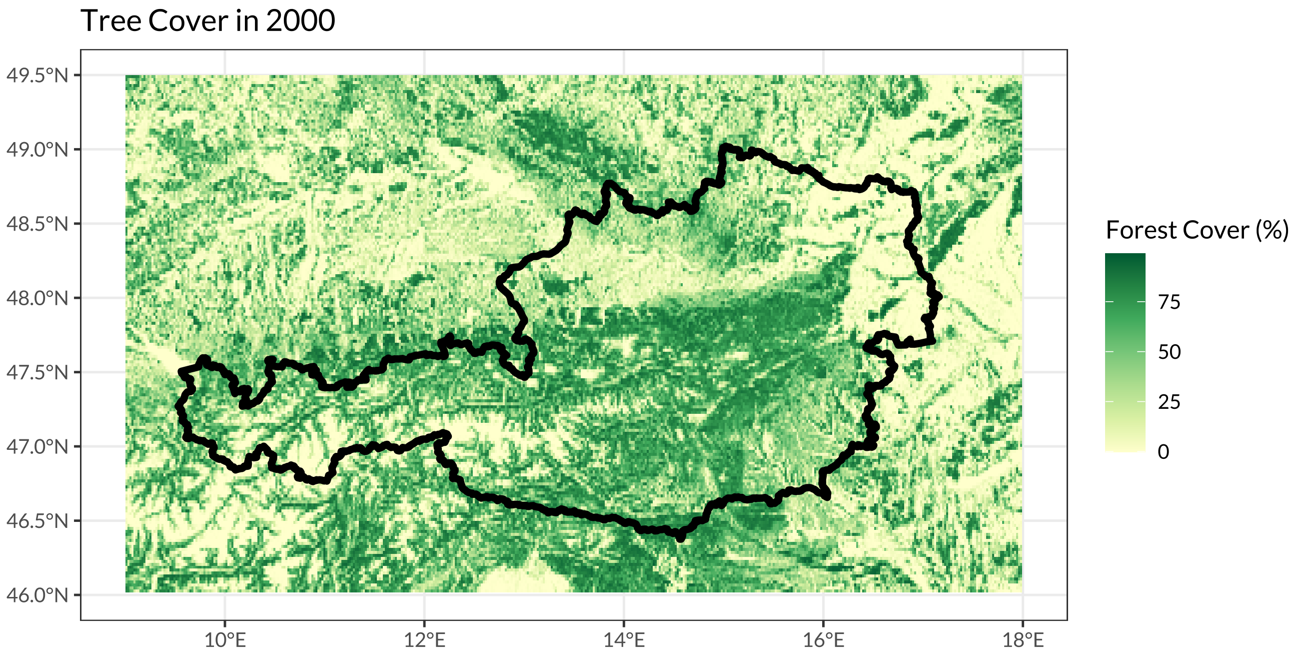
This visualization builds on Elia's and my try from the lecture, but is coded again from scratch since I cleverly decided to not save the script I used then. The main variable that is plotted as different fill colors of municipalities is the percentual change of forest area from 2003 to 2022. Blue values represent a forest increase, and red values represent forest loss. The more intense a color is, the more pronounced the forest gain/loss. You may notice that there are no numbers in the legend. This is because I used an ordered scale, meaning that heavier deforestation will always be redder, but it cannot be inferred how *much* more heavy deforestation was. This is to more clearly show differences between municipalities, which have either very similar deforestation figures in the region of –50% to –100%, or positive values. Had I used a simple continuous scale, most of the map would be assigned more or less one of two colors (or at least it would look like it). Overlaid are dots representing locations of slaughterhouses. The idea behind this map is to encourage the viewer to perform eyeball econometrics and come to the conclusion that slaughterhouses (1) cluster in (2) red areas. I think that this visualization is appropriate for this cause since it comes to mind as a natural way of evaluating point data against a continuous spatial variable (deforestation). Note that I spent five hours trying to include a basemap background layer, failed to do that using both tmap and ggplot, and then abandoned the idea out of time considerations and thought, "I should have used QGIS."

Task B

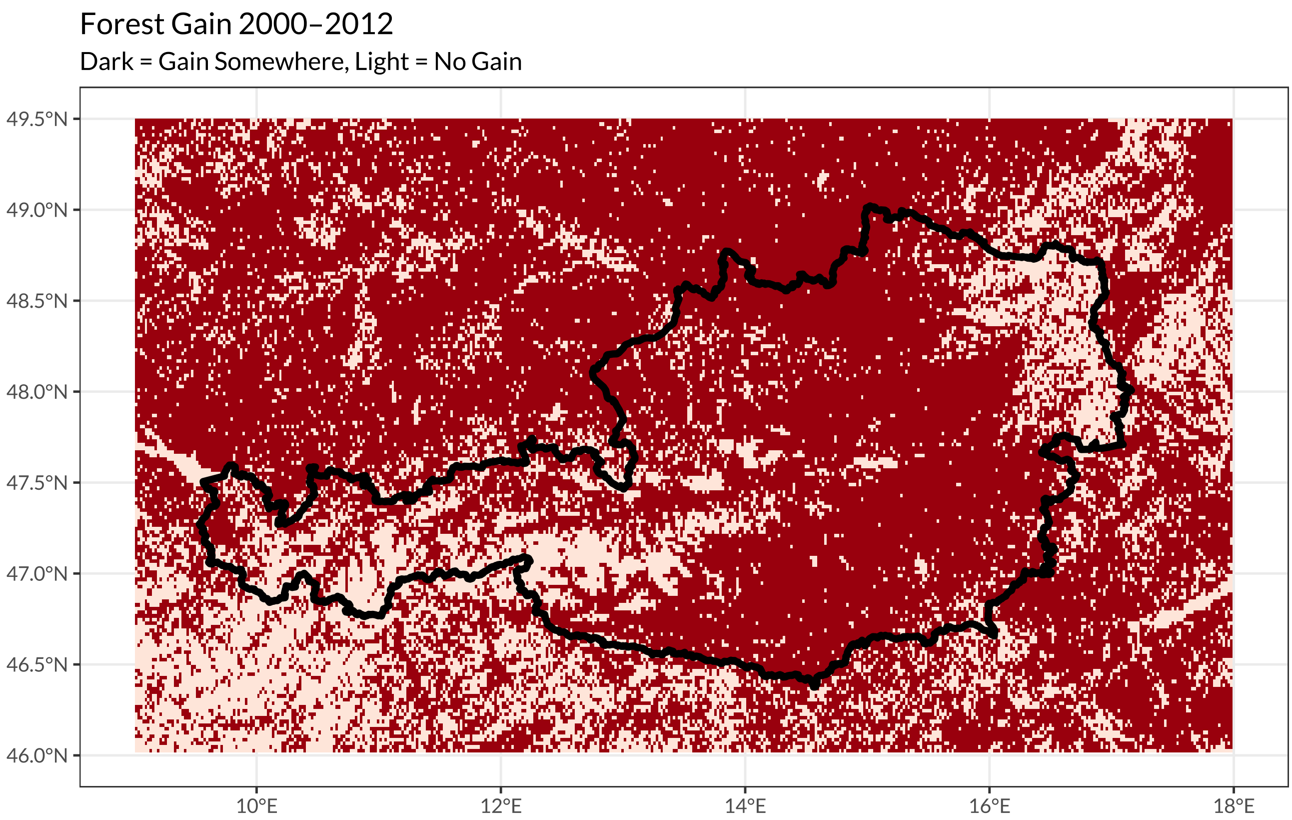
I begin by creating the following visualization of the "forest cover 2000" band from portions of two adjacent tiles. I aggregated the data quite radically, from a 1 arc second resolution to a 1 arc minute resolution, to make processing and plotting faster. It should suffice for these purposes. The following visualization shows percentage of forest cover in a given cell. This could be more insightful if we knew where this is.



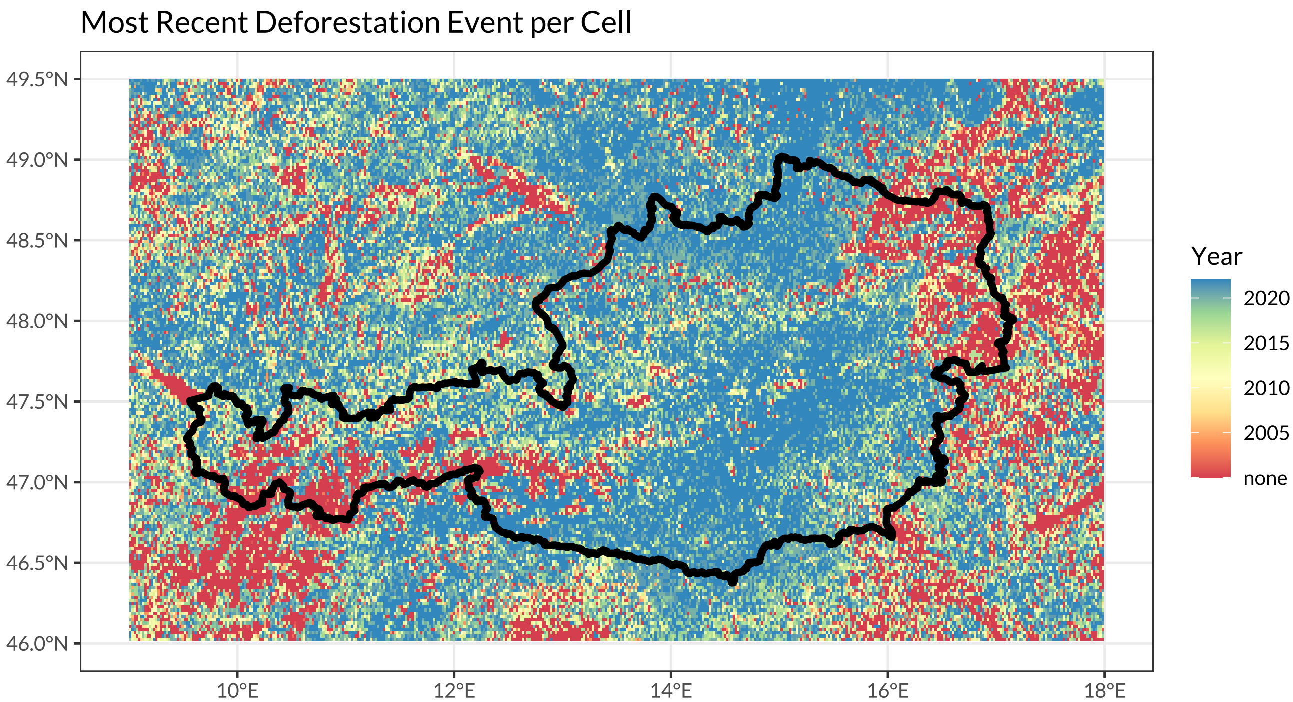
Thus, the following visualization adds the borders of Austria. We now know where we are and can more easily interpret the data. We can see that forest cover is very high in mountaineous areas, with the exception of regions at high altitude, and lowest in regions where many people live. This makes intuitive sense.



Because the provided data has more information than just forest cover in 2000, I created two other, maybe more insightful visualizations. In the first of these two, red cells represent areas where at least one original cell experienced forest *gain* between 2000 and 2012. We can see that most 1-arc-minute cells have at least one 1-arc-second cell in them that was transformed to forest within that time frame.

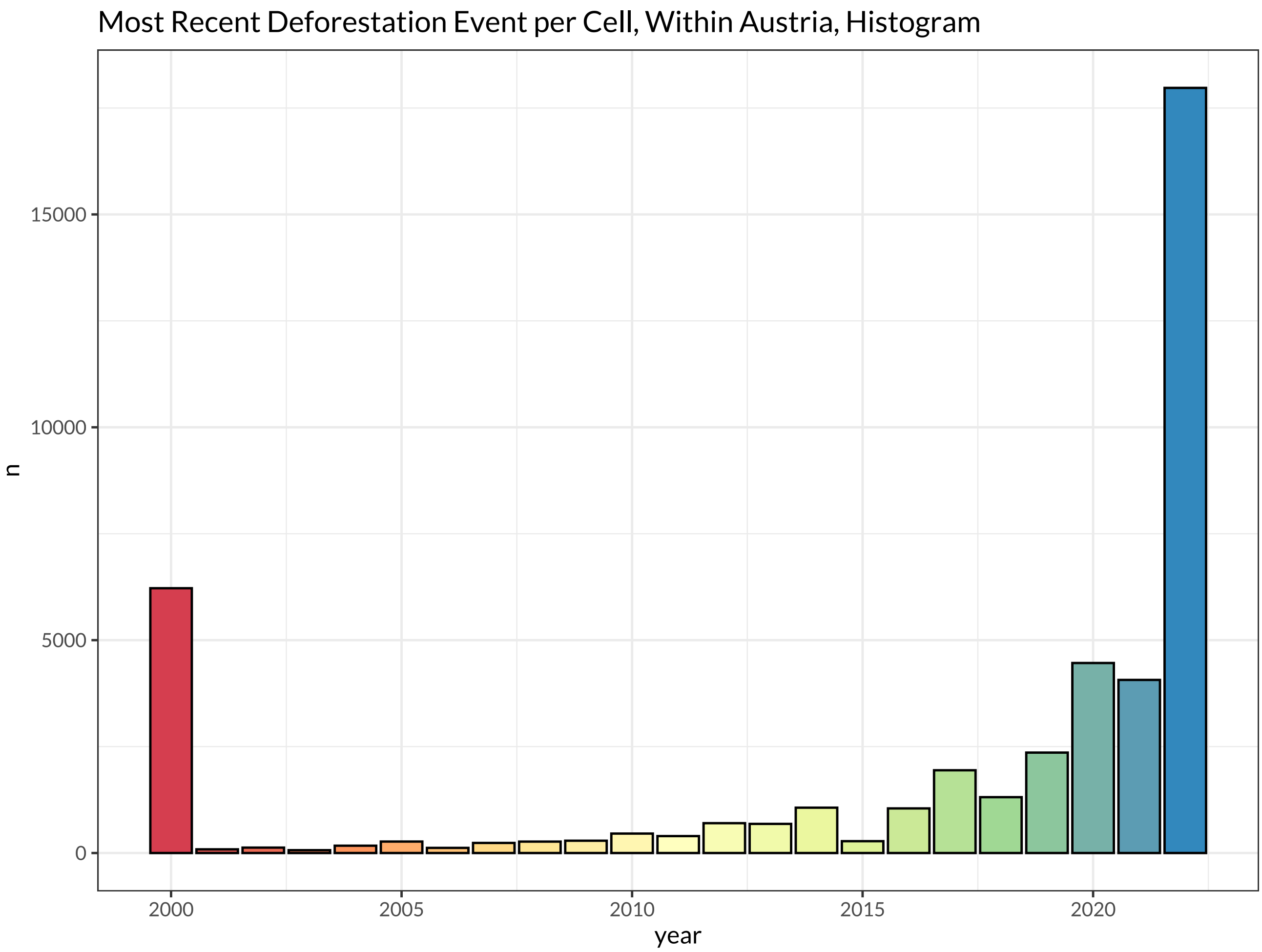


The second additional visualization shows when the most recent deforestation event for each cell happened. This means that if a cell is colored in the color representing "2020," then some pixel inside the aggregated cell experienced deforestation in 2020, and no pixel experienced deforestation later than 2020.



We can see that red cells, meaning those where no deforestation at all happened, cluster in southwestern Tyrol and northeastern Lower Austria. This coincides with areas where there was also no forest *gain*. It also coincides with low-forest areas from the original visualization. Thus, a likely explanation for why there was no deforestation at all in these cells is that there is just not too much happening in terms of forest there.

What if we now create a visualization without any spatial information? To discuss this, I subset the data to cells within Austria (as opposed to the above visualization) and plot a histogram of cell values. To facilitate interpretation, colors are kept the same as in the previous visualization.



While we *could guess* from the map that 2022 is the most common value, determining whether 2020 and 2021 are more common than “none” or not is difficult. This kind of information can be more easily read off the histogram. However, since I stripped the data of all its spatial information, every nuance of *where* is missing. We could not relate the count of “no deforestation” cells with any spatial explanations, such as terrain, population patterns, spatial autocorrelation with other no-deforestation cells, and so on.

Task C

I fear I do not completely understand the way the question is asked, I will attempt an answer in the following anyway.

First, I am not sure whether we should consider the *same* type of connections at every level of aggregation. I find very few variables that would make sense at any level of aggregation. One variable that would make sense is trade flows.

So consider the following networks: On the individual level, Alice trades with Bob, Bob trades with Alice, they may even trade with a person whose name starts with C, and so on. On the sub-state group level, towns may trade with one another. On the state level, there is trade happening too. As for the planet level, it is a known fact that the sum of all trade balances is positive, so there probably is trade with space too. The advantage here is that trade flows can be measured in monetary (or even real goods) terms on every level of aggregation.

On the individual level, I expect the network to be least dense. Every individual has only so much time, and they can only interact with so many people. Thus, the probability of there being a link between two randomly drawn people is probably low. On any more general level of aggregation, density may increase since the probability that *somebody* from New York is trading with *somebody* from Chicago is much higher than for any two individual people from these cities.

Regarding centrality, I expect there to be a much more even centrality distribution at higher levels of aggregation. At the individual level, I expect there to be very few agents who are very central, e.g., because they trade with a product that is in high demand. If we look at the other side of the aggregation scale, there are countries that trade more and countries that trade less, but the differences are probably smaller.

If links are trade links, I expect them to be influenced by personal ties between people, wealth, and regional proximity. If two people are good friends, I would expect them to trade more just because they interact more with another. I also would expect wealthy people to trade more since trade is most often conducted goods against money, and with more money, you can trade more goods. Regional proximity may play a role because increasing distance means increasing transportation costs.

On the individual level, I expect the personal ties mechanism to be much stronger. On higher levels of aggregation, I would expect the regional proximity channel to predominate, just because individual connections do not matter as much on these levels. Wealth is probably a factor on any level of aggregation.